世纪统计学系列教材

应用时间序列分析

王 燕 编著

中国人民大学出版社

VISIT…

LANZAROTE Caliente.com

21世纪统计学系列教材

应用时间序列分析

王　燕　编著

中国人民大学出版社

图书在版编目(CIP)数据

应用时间序列分析/王燕编著.
北京:中国人民大学出版社,2005
(21世纪统计学系列教材)
ISBN 7-300-06654-2

Ⅰ.应…
Ⅱ.王…
Ⅲ.时间序列分析-高等学校-教材
Ⅳ.O211.61

中国版本图书馆CIP数据核字(2005)第069800号

21世纪统计学系列教材
应用时间序列分析
王　燕　编著

出版发行	中国人民大学出版社		
社　址	北京中关村大街31号	邮政编码	100080
电　话	010－62511242(总编室)		010－62511239(出版部)
	010－82501766(邮购部)		010－62514148(门市部)
	010－62515195(发行公司)		010－62515275(盗版举报)
网　址	http://www.crup.com.cn		
	http://www.ttrnet.com(人大教研网)		
经　销	新华书店		
印　刷	北京雅艺彩印有限公司		
开　本	787×965毫米 1/16	版　次	2005年7月第1版
印　张	17.25	印　次	2005年7月第1次印刷
字　数	312 000	定　价	23.00元(含光盘)

版权所有　侵权必究　　印装差错　负责调换

《21世纪统计学系列教材》编委会

编委会主任 易丹辉

编委会委员 （按姓氏笔画排序）

尹德光 冯士雍 张尧庭

陈希孺 吴喜之 赵彦云

柯惠新 袁 卫 倪加勋

顾 岚 袁寿庄 耿 直

总　序

改革开放以来，高等统计教育有了很大的发展。随着课程设置的不断调整，有不少教材出版，同时也翻译引进了一些国外优秀教材。作为培养我国统计专门人才的摇篮，中国人民大学统计学系自 1952 年创建以来，走过了风风雨雨，一直坚持着理论与应用相结合的办学方向，培养能够理论联系实际、解决实际问题的高层次人才。随着新知识经济和网络时代的到来，我们在教学科研的实践中，深切地感受到，无论是自然科学领域、社会科学领域的研究，还是国家宏观管理和企业生产经营管理，甚至人们的日常生活，信息需求量日益增多，信息处理技术更加复杂，作为信息技术支柱的统计方法，越来越广泛地应用于各个领域。

面对新的形势，我们一直在思索，课程设置、教材选择、教学方式等怎样才能使学生适应社会经济发展的客观需要。在反复酝酿、不断尝试的基础上，我们决定与统计学界的同仁，共同编写、出版一套面向 21 世纪的统计学系列教材。

这套系列教材聘请了中国科学院院士、中国科技大学陈希孺教授，上海财经大学数量经济研究院张尧庭教授，中国科学院数学与系统科学研究所冯士雍研究员等作为编委。他们长期任中国人民大学的兼职教授，一直关心、支持着统计学系的学科建设和应用统计的发展。中国人民大学应用统计科学研究中心 2000 年已成为国家级研究基地，这些专家是首批专职或兼职研究人员。这一开放性研究

基地的运作，将有利于提升我国应用统计科学研究的水平，也必将进一步促进高等统计教育的发展。

这套教材是我们奉献给新世纪的，希望它能促进应用统计教育水平的提高。这套教材力求体现以下特点：

第一，在教材选择上，主要面向经济类统计学专业。选材既包括统计教材也包括风险管理与精算方面的教材。尽管名为统计学系列教材，但并不求大、求全，面是力求精选。对于目前已有的内容较为成熟、适合教学需要、公认的较好的教材，并未列入本次出版计划。

第二，每部教材的内容和写作，注意广泛吸收国内外优秀教材的成果。教材力求简明易懂、内容系统和实用，注重对统计方法思想的阐述，并结合大量实际数据和实例说明统计方法的特点及应用条件。

第三，强调与计算机的结合。为着力提高学生运用统计方法分析解决问题的能力，教材所涉及的统计计算，要求运用目前已有的统计软件。根据教材内容，选择使用 SAS，SPSS，TSP，STATISTICA，EViews，MINITAB，Excel 等。

感谢中国人民大学出版社的同志们，他们怀着发展我国应用统计科学的热情和据高统计教育水平的愿望，经过反复论证，使这套教材得以出版。感谢参与教材编写的同行专家、统计学系的教师。愿大家的辛勤劳动能够结出丰硕的果实。我们期待着与统计学界的同仁，共同创造应用统计辉煌的明天。

易丹辉

2000 年 8 月

于中国人民大学

前　言

所谓时间序列就是按照时间的顺序记录的一列有序数据。对时间序列进行观察、研究，找寻它变化发展的规律，预测它将来的走势就是时间序列分析。在日常生产、生活中，时间序列比比皆是，时间序列分析有着非常广泛的应用领域。

作为数理统计学的一个专业分支，时间序列分析遵循数理统计学的基本原理，都是利用观察信息估计总体的性质。但是由于时间的不可重复性，使得我们在任意一个时刻只能获得惟一的一个序列观察值，这种特殊的数据结构导致时间序列分析有它非常特殊的、自成体系的一套分析方法。

目前，国内有关时间序列分析的著作和教材有很多，每本书都会有它预定的读者群体。本书的定位是大学本科生的时间序列分析入门教材。根据这个定位，本书语言通俗、案例丰富，理论联系实际紧密，习题难易程度适当，非常便于学生理解和练习。

随着计算机科学的高速发展，现在有许多软件可以帮助我们进行时间序列分析。其中最权威的软件是SAS。在SAS系统中有一个专门进行计量经济与时间序列分析的模块：SAS/ETS (Econometric & Time Series)。同时，由于SAS系统具有全球一流的数据仓库功能，因此在进行海量数据的时间序列分析时具有其他统计软件无可比拟的优势。

为了帮助同学们在学习理论知识的同时也能熟练地掌握 SAS/ETS 软件的操作和分析技巧，本书在每一章后面都会有一小节的内容详细介绍本章的分析方法在 SAS 软件上的实现。为使同学们更好地学习和操作，本书所有例题的数据、习题数据、例题的操作程序及上机指导程序都制成电子版本，刻录成光盘随书附赠。为了便于教师上课，光盘中还附赠全书的 PowerPoint 课件。

编著者

2005 年 5 月

目　录

第1章　时间序列分析简介 …………………………………………………………… (1)

1.1　引言 ………………………………………………………………………… (1)

1.2　时间序列的定义 …………………………………………………………… (2)

1.3　时间序列分析方法 ………………………………………………………… (2)

1.3.1　描述性时序分析 ……………………………………………………… (2)

1.3.2　统计时序分析 ………………………………………………………… (3)

1.4　时间序列分析软件 ………………………………………………………… (5)

1.5　习题 ………………………………………………………………………… (6)

1.6　上机指导 …………………………………………………………………… (6)

1.6.1　SAS 操作界面 ………………………………………………………… (6)

1.6.2　创建时间序列 SAS 数据集 …………………………………………… (7)

1.6.3　时间序列数据集的处理 ……………………………………………… (11)

第2章　时间序列的预处理 ………………………………………………………… (16)

2.1　平稳性检验 ………………………………………………………………… (16)

2.1.1　特征统计量 …………………………………………………………… (16)

2.1.2　平稳时间序列的定义 ………………………………………………… (18)

2.1.3 平稳时间序列的统计性质…………………………………… (20)
2.1.4 平稳时间序列的意义…………………………………………… (21)
2.1.5 平稳性的检验…………………………………………………… (22)
2.2 纯随机性检验……………………………………………………… (27)
2.2.1 纯随机序列的定义……………………………………………… (28)
2.2.2 白噪声序列的性质……………………………………………… (28)
2.2.3 纯随机性检验…………………………………………………… (29)
2.3 习题………………………………………………………………… (34)
2.4 上机指导…………………………………………………………… (36)
2.4.1 绘制时序图……………………………………………………… (36)
2.4.2 平稳性与纯随机性检验………………………………………… (38)
第3章 平稳时间序列分析 ………………………………………… (41)
3.1 方法性工具………………………………………………………… (41)
3.1.1 差分运算………………………………………………………… (41)
3.1.2 延迟算子………………………………………………………… (42)
3.1.3 线性差分方程…………………………………………………… (43)
3.2 *ARMA* 模型的性质 …………………………………………… (44)
3.2.1 *AR* 模型 ……………………………………………………… (44)
3.2.2 *MA* 模型 ……………………………………………………… (59)
3.2.3 *ARMA* 模型 …………………………………………………… (65)
3.3 平稳序列建模……………………………………………………… (68)
3.3.1 建模步骤………………………………………………………… (68)
3.3.2 样本自相关系数与偏自相关系数……………………………… (69)
3.3.3 模型识别………………………………………………………… (70)
3.3.4 参数估计………………………………………………………… (74)
3.3.5 模型检验………………………………………………………… (80)
3.3.6 模型优化………………………………………………………… (82)
3.4 序列预测…………………………………………………………… (88)
3.4.1 线性预测函数…………………………………………………… (88)
3.4.2 预测方差最小原则……………………………………………… (89)
3.4.3 线性最小方差预测的性质……………………………………… (90)
3.4.4 修正预测………………………………………………………… (95)
3.5 习题………………………………………………………………… (97)

3.6 上机指导 …… (101)
3.6.1 模型识别 …… (102)
3.6.2 参数估计 …… (104)
3.6.3 序列预测 …… (106)
第4章 非平稳序列的确定性分析 …… (108)
4.1 时间序列的分解 …… (108)
4.1.1 Wold 分解定理 …… (108)
4.1.2 Cramer 分解定理 …… (109)
4.2 确定性因素分解 …… (110)
4.3 趋势分析 …… (111)
4.3.1 趋势拟合法 …… (111)
4.3.2 平滑法 …… (114)
4.4 季节效应分析 …… (119)
4.5 综合分析 …… (122)
4.6 X-11 过程 …… (127)
4.7 习题 …… (129)
4.8 上机指导 …… (131)
4.8.1 拟合线性趋势 …… (131)
4.8.2 拟合非线性趋势 …… (133)
4.8.3 X-11 过程 …… (136)
第5章 非平稳序列的随机分析 …… (139)
5.1 差分运算 …… (140)
5.1.1 差分运算的实质 …… (140)
5.1.2 差分方式的选择 …… (141)
5.1.3 过差分 …… (145)
5.2 *ARIMA* 模型 …… (146)
5.2.1 *ARIMA* 模型的结构 …… (146)
5.2.2 *ARIMA* 模型的性质 …… (147)
5.2.3 *ARIMA* 模型建模 …… (149)
5.2.4 *ARIMA* 模型预测 …… (152)
5.2.5 疏系数模型 …… (155)
5.2.6 季节模型 …… (159)
5.3 Auto-Regressive 模型 …… (167)

5.3.1 模型结构 …… (167)
5.3.2 残差自相关检验 …… (169)
5.3.3 模型拟合 …… (172)
5.4 异方差的性质 …… (175)
5.4.1 异方差的影响 …… (175)
5.4.2 异方差的直观诊断 …… (176)
5.5 方差齐性变换 …… (179)
5.6 条件异方差模型 …… (182)
5.6.1 模型结构 …… (182)
5.6.2 模型拟合 …… (185)
5.7 习题 …… (191)
5.8 上机指导 …… (193)
5.8.1 拟合 *ARIMA* 模型 …… (193)
5.8.2 拟合 *Auto-Regressive* 模型 …… (196)
5.8.3 拟合 *GARCH* 模型 …… (203)
第六章 多元时间序列分析 …… (208)
6.1 平稳多元序列建模 …… (209)
6.2 虚假回归 …… (212)
6.3 单位根检验 …… (213)
6.3.1 DF 检验 …… (214)
6.3.2 ADF 检验 …… (218)
6.3.3 PP 检验 …… (221)
6.4 协整 …… (225)
6.4.1 单整与协整 …… (225)
6.4.2 协整检验 …… (226)
6.5 误差修正模型 …… (229)
6.6 习题 …… (230)
6.7 上机指导 …… (232)
附录 1 …… (240)
附录 2 …… (256)
附录 3 …… (259)
参考文献 …… (262)

第 1 章

时间序列分析简介

1.1 引　言

最早的时间序列分析可以追溯到 7 000 年前的古埃及。当时，为了发展农业生产，古埃及人一直在密切关注尼罗河泛滥的规律。把尼罗河涨落的情况逐天记录下来，就构成所谓的时间序列。对这个时间序列长期的观察使他们发现尼罗河的涨落非常有规律。天狼星第一次和太阳同时升起的那一天之后，再过两百天左右，尼罗河就开始泛滥，泛滥期将持续七八十天，洪水过后，土地肥沃，随意播种就会有丰厚的收成。由于掌握了尼罗河泛滥的规律，古埃及的农业迅速发展，解放出大批的劳动力去从事非农业生产，从而创建了古埃及灿烂的史前文明。

像古埃及人一样，按照时间的顺序把随机事件变化发展的过程记录下来就构成了一个时间序列。对时间序列进行观察、研究，找寻它变化发展的规律，预测它将来的走势就是时间序列分析。

1.2 时间序列的定义

在统计研究中，常用按时间顺序排列的一组随机变量

$$\cdots, X_1, X_2, \cdots, X_t, \cdots \tag{1.1}$$

来表示一个随机事件的时间序列，简记为 $\{X_t, t\in T\}$ 或 $\{X_t\}$。

用

$$x_1, x_2, \cdots, x_n \tag{1.2}$$

或 $\{x_t, t=1, 2, \cdots, n\}$ 表示该随机序列的 n 个有序观察值，称之为序列长度为 n 的观察值序列，有时也称式（1.2）为式（1.1）的一个实现。

在日常生产生活中，观察值序列比比皆是。比如把北京市城镇居民 1990—1999 年每年的消费支出按照时间顺序记录下来，就构成了一个序列长度为 10 的消费支出时间序列（单位：亿元）：

1 686，1 925，2 356，3 027，3 891，4 874，5 430，5 796，6 217，6 796

我们研究的目的是想揭示随机时序 $\{X_t\}$ 的性质，而要实现这个目标就是通过分析它的观察值序列 $\{x_t\}$ 的性质，由观察值序列的性质来推断随机时序 $\{X_t\}$ 的性质。

1.3 时间序列分析方法

1.3.1 描述性时序分析

早期的时序分析通常都是通过直观的数据比较或绘图观测，寻找序列中蕴含的发展规律，这种分析方法就称为描述性时序分析。古埃及人发现尼罗河泛滥的规律就是依靠这种分析方法。而在天文、物理、海洋学等自然科学领域，这种简单的描述性时序分析方法也常常能使人们发现意想不到的规律。

比如，19 世纪中后叶，德国药剂师、业余天文学家施瓦尔就是运用这种方法，经过几十年不断的观察、记录，发现了太阳黑子的活动具有 11 年左右的周期（数据见附录 1.1），如图 1—1 所示。

描述性时序分析方法具有操作简单、直观有效的特点，从史前直到现在一直被人们广为使用，它通常是人们进行统计时序分析的第一步。

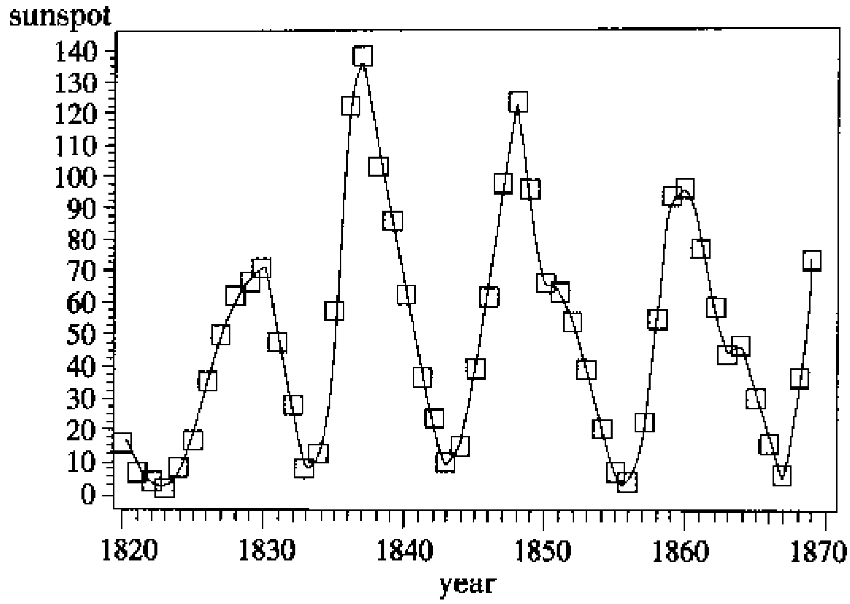


图 1—1　1820—1869 年太阳黑子年度数据时序图

1.3.2　统计时序分析

随着研究领域的不断拓广，人们发现单纯的描述性时序分析有很大的局限性。在金融、保险、法律、人口、心理学等社会科学研究领域，随机变量的发展通常会呈现出非常强的随机性，想通过对序列简单的观察和描述，总结出随机变量发展变化的规律，并准确预测出它们将来的走势通常是非常困难的。

为了更准确地估计随机序列发展变化的规律，从 20 世纪 20 年代开始，学术界利用数理统计学原理分析时间序列。研究的重心从表面现象的总结转移到分析序列值内在的相关关系上，由此开辟了一门应用统计学科——时间序列分析。

纵观时间序列分析方法的发展历史可以将时间序列分析方法分为两大类。

一、频域（frequency domain）分析方法

频域分析方法也被称为“频谱分析”或“谱分析”（spectral analysis）方法。

早期的频域分析方法假设任何一种无趋势的时间序列都可以分解成若干不同频率的周期波动，借助富里埃分析从频率的角度揭示时间序列的规律，后来又借助了傅里叶变换，用正弦、余弦项之和来逼近某个函数。20 世纪 60 年代，Burg 在分析地震信号时提出最大熵谱估计理论，该理论克服了传统谱分析所固有的分辨率不高和频率漏泄等缺点，值谱分析进入一个新阶段，我们称之为现代谱分析阶段。

目前谱分析方法主要运用于电力工程、信息工程、物理学、天文学、海洋学

和气象科学等领域，它是一种非常有用的纵向数据分析方法。但是由于谱分析过程一般都比较复杂，研究人员通常要具有很强的数学基础才能熟练使用它，同时它的分析结果也比较抽象，不易于进行直观解释，导致谱分析方法的使用具有很大的局限性。

二、时域（time domain）分析方法

时域分析方法主要是从序列自相关的角度揭示时间序列的发展规律。相对于谱分析方法，它具有理论基础扎实、操作步骤规范、分析结果易于解释的优点。目前它已广泛应用于自然科学和社会科学的各个领域，成为时间序列分析的主流方法。我们这本书主要就是介绍时域分析方法。

时域分析方法的基本思想是源于事件的发展通常都具有一定的惯性，这种惯性用统计的语言来描述就是序列值之间存在着一定的相关关系，而且这种相关关系具有某种统计规律。我们分析的重点就是寻找这种规律，并拟合出适当的数学模型来描述这种规律，进而利用这个拟合模型来预测序列未来的走势。

时域分析方法具有相对固定的分析套路，通常都遵循如下分析步骤：

第一步：考察观察值序列的特征。

第二步：根据序列的特征选择适当的拟合模型。

第三步：根据序列的观察数据确定模型的口径。

第四步：检验模型，优化模型。

第五步：利用拟合好的模型来推断序列其他的统计性质或预测序列将来的发展。

时域分析方法的产生最早可以追溯到 1927 年，英国统计学家 G. U. Yule（1871—1951）提出自回归（autoregressive，AR）模型。不久之后，英国数学家、天文学家 G. T. Walker 爵士在分析印度大气规律时使用了移动平均（moving average ，MA）模型和自回归移动平均（autoregressive moving average，ARMA）模型（1931 年）。这些模型奠定了时间序列时域分析方法的基础。

1970 年，美国统计学家 G. E. P. Box 和英国统计学家 G. M. Jenkins 联合出版了 *Time Series Analysis Forecasting and Control* 一书。在书中，Box 和 Jenkins 在总结前人研究的基础上，系统地阐述了对求和自回归移动平均（autoregressive integrated moving average ，ARIMA）模型的识别、估计、检验及预测的原理及方法。这些知识现在被称为经典时间序列分析方法，是时域分析方法的核心内容。为了纪念 Box 和 Jenkins 对时间序列发展的特殊贡献，现在人们也常把 ARIMA 模型称为 Box-Jenkins 模型。

Box-Jenkins 模型实际上是主要运用于单变量、同方差场合的线性模型。随

着人们对各领域时间序列的深入研究，发现该经典模型在理论和应用上都还存在着许多局限性。所以近 20 年来，统计学家纷纷转向多变量场合、异方差场合和非线性场合的时间序列分析方法的研究，并取得了突破性的进展。

在异方差场合，美国统计学家、计量经济学家 Robert F. Engle 在 1982 年提出了自回归条件异方差（ARCH）模型，用以研究英国通货膨胀率的建模问题。为了进一步放宽 ARCH 模型的约束条件，Bollerslov 在 1985 年提出了广义自回归条件异方差（GARCH）模型。随后 Nelson 等人又提出了指数广义自回归条件异方差模型（EGARCH）、方差无穷广义自回归条件异方差模型（IGARCH）和依均值广义自回归条件异方差（GARCH-M）等限制条件更为宽松的异方差模型。这些异方差模型是对经典的 ARIMA 模型很好的补充。它比传统的方差齐性模型更准确地刻画了金融市场风险的变化过程，因此 ARCH 模型及其衍生出的一系列拓展模型在计量经济学领域有着广泛的应用。Engle 也因此获得 2003 年诺贝尔经济学奖。

在多变量场合，Box 和 Jenkins 在 *Time Series Analysis Forecasting and Control* 一书中研究过平稳多变量序列的建模，Box 和 Tiao 在 1970 年左右讨论过带干扰变量的时间序列分析。这些研究实际上是把对随机事件的横向研究和纵向研究有机地融合在一起，提高了对随机事件分析和预测的精度。1987 年，英国统计学家、计量经济学家 C. Granger 提出了协整（co-integration）理论，进一步为多变量时间序列建模松绑。有了协整的概念之后，在多变量时间序列建模过程中“变量是平稳的”不再是必须条件了，而只要求它们的某种线性组合平稳。协整概念的提出极大地促进了多变量时间序列分析方法的发展，Granger 因此与 Engle 一起获得了 2003 年诺贝尔经济学奖。

非线性时间序列分析也有重大发展，汤家豪教授等在 1980 年左右提出了利用分段线性化构造的门限自回归模型成为目前分析非线性时间序列的经典模型。

1.4 时间序列分析软件

随着计算机科学的高速发展，现在有许多软件可以帮助我们进行时间序列分析。常用的软件有：S-plus，Matlab，Gauss，TSP，Eviews 和 SAS。

我们在本书中介绍和使用的软件是 SAS。SAS 的全称是 Statistical Analysis System，直译过来就是统计分析系统。它最早是由美国北卡罗来纳州立大学（North Carolina State University）的两位教授（A. J. Barr & J. H. Goodnight）

联合开发的，专门进行数学建模和统计分析的软件。发展到今天，它不仅成为统计分析领域的国际标准软件，而且已经成为具有完备的数据访问、数据管理、数据分析和数据呈现功能的大型集成化软件系统。由于领先的技术和全面的功能，它已经成为全球数据分析方面的首选软件。

在 SAS 系统中有一个专门的模块：SAS/ETS（Econometric & Time Series)，这是一个专门进行计量经济与时间序列分析的软件。SAS/ETS 编程语言简洁，输出功能强大，分析结果精确，是进行时间序列分析与预测的理想的软件。由于 SAS 系统具有全球一流的数据仓库功能，因此在进行海量数据的时间序列分析时它具有其他统计软件无可比拟的优势。

为了让读者在了解时间序列分析基本原理的同时也能掌握一定的实际操作技巧，本书在每一章后面都会有一小节的内容介绍本章分析方法在 SAS/ETS 软件上的实现。本书所有的例题也都是以 SAS/ETS 作为操作软件。

1.5 习　题

1.1　什么是时间序列？请收集几个生活中的观察值序列。

1.2　时域方法的特点是什么？

1.3　时域方法的发展轨迹是怎样的？

1.4　在附录 1 中选择几个感兴趣的序列，创建数据集。

1.6 上机指导

1.6.1　SAS 操作界面

在 Windows 操作系统下双击“The SAS System”的图标，立刻就启动了 SAS 系统，进入 SAS 操作界面。

SAS 的操作界面主要由三个部分组成：菜单栏、工具栏和窗口。见图 1—2。

一、菜单栏

菜单栏包括八个选项，它们分别是：

【File】——文件选项。它主要用于处理文件的打开、关闭、保存、输入、输出、打印及发送等任务。

【Edit】——编辑选项。它主要负责文件的复原、剪切、复制、粘贴、清除、



图 1—2　SAS 8.0 版本操作界面

选择、查找、替代等编辑工作。

【View】——视图选项。它主要用于切换不同的视窗。

【Tools】——工具选项。它主要提供各种编辑工具，并可以进行系统信息的更改与管理。

【Run】——程序运行选项。它负责提交程序给 SAS 运行。该选项只在当前主窗口为编辑窗口时出现。

【Solutions】——分析方案选项。它提供各种分析方法。

【Window】——窗口选项。它主要控制视窗的大小、排列方式及视窗间的切换。

【Help】——帮助选项。它负责提供各种帮助。

二、工具栏

菜单栏下是一列图形化的工具栏，提供了最为常用的一些功能。根据图 1—2 的工具栏从左到右排列，它们分别是：命令输入及提交执行窗口、创建新文件、打开旧文件、保存、打印、预览、剪切、复制、粘贴、复原、目录库、资源管理器、提交程序给 SAS 运行、删除、中断运行、帮助等 16 个工具选项。

三、窗口

SAS 有三个基本窗口：程序编辑窗口（Program Editor），运行记录窗口（Log）和结果输出窗口（Output）。我们将在程序编辑窗口编辑、修改程序；在运行记录窗口观察程序运行状况，获得错误提示；在结果输出窗口观看程序运行结果。

1.6.2　创建时间序列 SAS 数据集

要使用 SAS 分析数据，首先要将数据转换成 SAS 系统能够分析的 SAS 数据集（SAS datasets）。

根据数据的不同来源，SAS 系统提供了多种创建 SAS 数据集的方法，我们以如下数据为例，介绍两种最常用的创建 SAS 数据集的办法。见表 1—1。

表 1—1

时间	价格
2005 年 1 月	101
2005 年 2 月	82
2005 年 3 月	66
2005 年 4 月	35
2005 年 5 月	31
2005 年 6 月	7

一、使用 DATA 步创建 SAS 数据集

1. 创建临时数据集

在程序编辑窗口（Program Editor ）输入如下命令，即可产生一个名字为 example1 _ 1 的临时数据集储存上述数据。

```
data example1 _ 1;
input time monyy7. price;
format time monyy5. ;
cards;
Jan2005    101
Feb2005    82
Mar2005    66
Apr2005    35
May2005    31
Jun2005    7
;
run;
```

语句说明：

(1) SAS 系统的命令语句不分字母的大小写，单词之间至少空一格，每条命令都用分号"；"作为结束提示。

(2) "data example1 _ 1;"

这是命令 SAS 系统建立一个名字为 example1 _ 1 的临时数据集。该数据集在本次 SAS 系统运行过程中会始终保存在一个名为 WORK 的数据库中，一旦退出本次 SAS 系统运行，该数据集将不存盘退出。

(3) "input time monyy7. price;"

这是告诉 SAS 系统，我们要输入两个变量的数据。

第一个变量取名为“time”，“monyy7.”，说明该变量是时间变量，且指定了数据的输入格式为字符长度为 7 的月份—年度数据，输入格式为月份的 3 位缩写字母加 4 位年份数据，形如“Jan2005”。

第二个变量取名为“price”。对它我们没有指定变量类型和数据输入格式，那么系统就自动将它视为数值型变量，自动读取任意长度的变量值。

对不同类型的数据，SAS 支持多种类型的输入、输出格式，常用输入、输出格式见附录 2。

（4）“format time monyy5.;”

这是告诉 SAS 系统，“time”这个变量的输出格式是字符长度为 5 的月份—年度数据，输出格式为月份的 3 位缩写字母加 2 位年份数据，形如“Jan05”。否则 SAS 会按照它的系统时间格式输出“time”的值。

（5）“cards;”

这是告诉 SAS 系统，下面开始输入数据行了。接着就是数据录入。我们既可以通过光标引导键入数据，也可以通过复制、粘贴等功能从其他文件中将数据复制过来。

在此数据以列的方式被读取，第一列数据会自动赋值给变量“time”，第二列数据会自动赋值给变量“price”，其余的数据将被忽略。

如果将第二句命令修改为“input time monyy7. price@@;”，则数据将会以行的方式被读取：第一个数据赋值给变量“time”，第二个数据赋值给变量“price”，第三个数据赋值给变量“time”，第四个数据赋值给变量“price”……

数据输入完毕后，另起一行输入命令结束符号“;”。

（6）“run;”

这是告诉系统程序写好了，可以运行了。

点击工具栏上 SUBMIT 图标（■），提交这个程序开始运行，运行结束后一个名为 example1 _ 1 的临时数据集就建立了，在本次 SAS 运行中，我们可以随时调用这个数据集。

2. 创建永久数据集

SAS 规定每个数据集使用二级命名制——数据库名 . 数据集名。所谓“数据库”类似于文件夹的概念，是一个专门存储数据集的仓库。SAS 系统本身提供了两个通用数据库。

一个是存放临时数据集的数据库——WORK 数据库。该库中的数据集在本次 SAS 系统运行中始终存在，而一旦退出本次操作，该库中的数据集将被

全部删除。SAS 会自动地把所有省缺数据库名的数据集存于 WORK 数据库中。

另一个是永久数据库——SASUSER 数据库。所谓永久数据库就是指在该库建立的数据集不会因为你退出 SAS 系统而丢失，它会永久地保存在该数据库中，你以后进入 SAS 系统还可以从该库中调用该数据集。

使用如下命令就可以得到一个名字为 sasuser. example1 _ 1 的永久数据集。

data sasuser. example1 _ 1;

除了“SASUSER”这个永久数据库之外，我们还可以使用 libname 命令建立自己的永久数据库。

假定我们的 SAS 软件安装在 D 盘，并且已经在 SAS 目录下建立一个名为 myfile 的子目录（文件夹），就可以通过如下命令在该目录下建立一个名为 datafile 的永久数据库，并将数据集 example1 _ 1 存入这个永久库中。

Libname datafile ‘d：\ sas \ myfile’;

data datafile. example1 _ 1;

以后这个数据集将一直以 datafile. example1 _ 1 这种二级名称的形式被引用。

3. 查看数据集

建立好一个数据集之后，我们可以使用 PRINT 程序查看这个数据集的结果。命令格式为：

proc print data=数据库名．数据集名；

在原程序中添加如下命令可以查看临时数据集 example1 _ 1 的内容：

proc print data=example1 _ 1;

run;

运行该程序，在结果输出窗口（Output），得到如下输出结果：

Obs	time	price
1	JAN05	101
2	FEB05	82
3	MAR05	66
4	APR05	35
5	MAY05	31
6	JUN05	7

假如从输出窗口发现数据集的数据输入有误，可以返回程序编辑窗口，修改程序及数据。

二、直接导入外部数据文件转换成 SAS 数据集

SAS 系统不仅允许我们在系统内部通过程序编辑的方式创建数据集，而且还允许我们将多种形式的外部数据文件直接导入，转换成 SAS 数据集。

假定我们先将上述数据存为 Excel 文件，文件名为 example1 _ 1. xls。通过如下操作，这个 Excel 文件可以直接转换成 SAS 数据集。

第一步：选择导入外部数据文件选项。

在菜单栏中，点击【File】选项，拉下文件管理菜单，点击其中的输入数据选项（Import Data)。

第二步：选择要输入数据的类型。

进入输入数据界面之后，首先是要选择输入数据的类型，SAS 的默认设置是 Microsoft Excel 97 or 2000（*. xls)，我们的文件正好是这个格式，选择 Next 选项，进行下一步。

第三步：指明该输入文件的路径。

点中 BROWSE 选项，指明输入文件 example1 _ 1. xls 的路径，选择 Next 选项，进行下一步。

第四步：指定该文件转换成 SAS 数据集后存放的数据库及数据集名。假定我们指定的是 SASUSER 库，文件名取为 example1 _ 1，点击 Finish 选项，数据格式转换完成之后，一个名为 sasuser. example1 _ 1 的永久 SAS 数据集就建立好了，我们直接调用该数据集进行分析。

1.6.3 时间序列数据集的处理

一、间隔函数的使用

如果没有现成的电子版数据导入，而是需要通过键盘录入数据，那么时间数据的录入将是非常麻烦的一件事情。对于等时间间隔数据，SAS 提供了一种间隔函数 INTNX，它可以根据需要自动产生等时间间隔的时间数据。

例如运行如下程序：

```
data example1 _ 2;
input price ;
time=intnx ( 'month', '01jan2005' d, _n_-1);
format time monyy. ;
cards;
  3.41
  3.45
```

```
    3.42
    3.53
    3.45
;
proc print data=example1 _ 2 ;
run;
```

得到输出结果如下：

Obs	price	time
1	3.41	JAN05
2	3.45	FEB05
3	3.42	MAR05
4	3.53	APR05
5	3.45	MAY05

尽管我们没有输入时间变量的取值，但是 INTNX 函数会自动给每条输入数据进行时间变量赋值。

语句说明：

“time=intnx（‘month’，‘01jan2005’ d， _ n _ -1）；”

该命令是指定用 INTNX 函数给时间变量 TIME 赋值。具体操作是以 2005 年 1 月 1 日为起始时间，以月为时间间隔，从起始时刻 2005 年 1 月 1 日开始每读入一个 PRICE 的数据，就自动产生一个 TIME 的时间数据。

INTNX 函数包括三个参数：

第一个参数是指定等时间间隔，本例中指定等时间间隔为月‘month’，该参数还可以取为 day（天），week（星期），quarter（季度），year（年）等。

第二个参数是指定参照时间，本例的参照时间是‘01jan2005’ d。

第三个参数是 _ n _ k，这个参数主要是调整开始观测指针。k 为整数，k 取正值，指针由参照时间向未来（不包含参照时间）拨 k 期，k 取负值，指针由参照时间向过去（包括参照时间）拨 k 期。

二、序列变换

在时间序列分析中，我们得到的是观察值序列 $\{x_t\}$，但是需要分析的可能却是这个观察值序列的某个函数变换。例如对数序列 $\{\ln x_t\}$。在建立数据集时，我们可以通过简单的赋值命令实现这个变换。

例如运行如下程序：

```
data example1 _ 3;
```

```
input price ;
logprice=log (price);
time=intnx ( 'month', '01jan2005' d,  _n_-1);
format time monyy. ;
cards;
  3.41
  3.45
  3.42
  3.53
  3.45
;
proc print data=example1 _ 3 ;
run;
```

该程序输出结果如下：

Obs	price	logprice	time
1	3.41	1.22671	JAN05
2	3.45	1.23837	FEB05
3	3.42	1.22964	MAR05
4	3.53	1.26130	APR05
5	3.45	1.23837	MAY05

语句说明：

"logprice=log (price);"

这是一个简单的赋值语句，将 price 的对数函数值赋值给一个新的变量 logprice，即建立了一个新的对数序列。

三、子集

有时我们只需要分析一个时间序列中的部分序列值，这时可以在 DATA 步中建立一个子集。

例如我们只需要分析数据集 example1 _ 3 中 2005 年 3 月～5 月的对数价格序列，建立相应的子集 example1 _ 4。

```
data example1 _ 4;
set example1 _ 3;
keep time logprice;
where time>='01mar2005'd;
```

```
proc print data=example1 _ 4;
run;
```

数据集 example1 _ 4 的输出结果如下：

Obs	logprice	time
1	1.22964	MAR05
2	1.26130	APR05
3	1.23837	MAY05

语句说明：

(1) "data example1 _ 4;"

这条命令是告诉系统要建立一个名为 example1 _ 4 的临时数据集。

(2) "set example1 _ 3;"

这条命令是说，数据集 example1 _ 4 是数据集 example1 _ 3 的子集。注意 example1 _ 3 数据集要在本次 SAS 运行中已经存在。

(3) "keep time logprice;"

这条命令是告诉系统数据集 example1 _ 4 中只需要保留两个变量：一个是 time；一个是 logprice。

(4) "where time>='01mar2005'd;"

这是指令系统只要将数据集 example1 _ 3 中时间大于等于 2005 年 3 月 1 日的那些数据输入数据集 example1 _ 4。

四、缺失值插值

有时我们的观察值序列会有缺失值，这会影响我们的分析。这时可以值用 EXPAND 过程，使用差值方法补全缺失值。

假设上例中 2005 年 3 月 1 日 price 的观察值缺失，使用如下程序对缺失值进行补插。

```
data example1 _ 5;
input price ;
time=intnx ( 'month', '01jan2005' d,  _ n _ -1);
format time date. ;
cards;
  3.41
  3.45

  3.53
```

```
3.45
;
proc expand data=example1 _ 5 out=example1 _ 6;
id time;
proc print data=example1 _ 5;
proc print data=example1 _ 6;
run;
```

该程序输出两个数据集的结果如下：

数据集 example1 _ 5

Obs	price	time
1	3.41	01JAN05
2	3.45	01FEB05
3	.	01MAR05
4	3.53	01APR05
5	3.45	01MAY05

数据集 example1 _ 6

Obs	time	price
1	01JAN05	3.41000
2	01FEB05	3.45000
3	01MAR05	3.50693
4	01APR05	3.53000
5	01MAY05	3.45000

语句说明：

“proc expand data=example1 _ 5 out=example1 _ 6;”

这是指令系统将数据集 example1 _ 5 中的所有缺失值用插值的方法补齐，并将补齐后的数据集另存为数据集 example1 _ 6。

第 2 章

时间序列的预处理

拿到一个观察值序列之后，首先要对它的平稳性和纯随机性进行检验，这两个重要的检验我们称之为序列的预处理。根据检验的结果可以将序列分为不同的类型，对不同类型的序列我们会采用不同的分析方法。

2.1 平稳性检验

2.1.1 特征统计量

平稳性是某些时间序列具有的一种统计特征。要描述清楚这个特征，我们必须借助如下的统计工具。

一、概率分布

数理统计的基础知识告诉我们分布函数或密度函数能够完整地描述一个随机变量的统计特征。同样，一个随机变量族 $\{X_t\}$ 的统计特性也完全由它们的联合分布函数或联合密度函数决定。

对于时间序列 $\{X_t, t\in T\}$，这样来定义它的概率分布：

任取正整数 m，任取 t_1，t_2，…，$t_m\in T$，则 m 维随机向量（X_{t_1}，X_{t_2}，…，X_{t_m})′的联合概率分布记为 $F_{t_1,t_2,\cdots,t_m}$（x_1，x_2，…，x_m)，由这些有限维分布函数构成的全体

$$\{F_{t_1,t_2,\cdots,t_m}(x_1,x_2,\cdots,x_m),\forall m\in(1,2,\cdots,m),\forall t_1,t_2,\cdots,t_m\in T\}$$

就称为序列 $\{X_t\}$ 的概率分布族。

概率分布族是极其重要的统计特征描述工具，因为序列的所有统计性质理论上都可以通过概率分布推测出来，但是概率分布族的重要性也就停留在这样的理论意义上。在实际应用中，要得到序列的联合概率分布几乎是不可能的，而且联合概率分布通常涉及非常复杂的数学运算，这些原因使我们很少直接使用联合概率分布进行时间序列分析。

二、特征统计量

一个更简单、更实用的描述时间序列统计特征的方法是研究该序列的低阶矩，特别是均值、方差、自协方差和自相关系数，它们也被称为特征统计量。

尽管这些特征量不能描述随机序列全部的统计性质，但由于它们概率意义明显，易于计算，而且往往能代表随机序列的主要概率特征，所以我们对时间序列进行分析，主要就是通过分析这些特征量的统计特性，推断出随机序列的性质。

1. 均使

对时间序列 $\{X_t, t\in T\}$ 而言，任意时刻的序列使 X_t 都是一个随机变量，都有它自己的概率分布，不妨记 X_t 的分布函数为 $F_t(x)$。只要满足条件 $\int_{-\infty}^{\infty} x\mathrm{d}F_t(x)<\infty$，就一定存在着某个常数 μ_t，使得随机变量 X_t 总是围绕在常数使 μ_t 附近做随机波动。我们称 μ_t 为序列 $\{X_t\}$ 在 t 时刻的均值函数。

$$\mu_t = EX_t = \int_{-\infty}^{\infty} x\mathrm{d}F_t(x)$$

当 t 取遍所有的观察时刻时，我们就得到一个均值函数序列 $\{\mu_t, t\in T\}$。它反映的是时间序列 $\{X_t, t\in T\}$ 每时每刻的平均水平。

2. 方差

当 $\int_{\infty}^{\infty} x^2\mathrm{d}F_t(x)<\infty$ 时，我们可以定义时间序列的方差函数用以描述序列值围绕其均值做随机波动时平均的波动程度。

$$DX_t = E(X_t-\mu_t)^2 = \int_{-\infty}^{\infty}(x-\mu_t)^2\mathrm{d}F_t(x)$$

同样，当 t 取遍所有的观察时刻时，我们得到一个方差函数序列 $\{Dx_t, t\in T\}$。

3. 自协方差函数和自相关函数

类似于协方差函数和相关系数的定义，在时间序列分析中我们定义自协方差函数（autocovariance function）和自相关系数（autocorrelation function）的概念。

对于时间序列 $\{X_t, t\in T\}$，任取 t，$s\in T$，定义 $\gamma(t,s)$为序列 $\{X_t\}$ 的自协方差函数：

$$\gamma(t,s)=E(X_t-\mu_t)(X_s-\mu_s)$$

定义 $\rho(t, s)$为时间序列 $\{X_t\}$ 的自相关系数，简记为 ACF。

$$\rho(t,s)=\frac{\gamma(t,s)}{\sqrt{DX_t \cdot DX_s}}$$

之所以称它们为自协方差函数和自相关系数是因为通常的协方差函数和相关系数度量的是两个不同的事件彼此之间的相互影响程度，而自协方差函数和自相关系数度量的是同一事件在两个不同时期之间的相关程度，形象地讲就是度量自己过去的行为对自己现在的影响。

2.1.2　平稳时间序列的定义

平稳时间序列有两种定义，根据限制条件的严格程度，分为严平稳时间序列和宽平稳时间序列。

一、严平稳（strictly stationary）

所谓严平稳就是一种条件比较苛刻的平稳性定义，它认为只有当序列所有的统计性质都不会随着时间的推移而发生变化时，该序列才能被认为平稳。而我们知道，随机变量族的统计性质完全由它们的联合概率分布族决定。所以严平稳时间序列的定义如下：

定义 2.1　设 $\{X_t\}$ 为一时间序列，对任意正整数 m，任取 t_1，t_2，…，$t_m\in T$，对任意整数 τ，有

$$F_{t_1,t_2\cdots t_m}(x_1,x_2,\cdots,x_m)=F_{t_{1+\tau},t_{2+\tau}\cdots t_{m+\tau}}(x_1,x_2,\cdots,x_m)$$

则称时间序列 $\{X_t\}$ 为严平稳时间序列。

前而说过，在实践中要获得随机序列的联合分布是一件非常困难的事，而且，即使知道随机序列的联合分布，计算和应用也非常不便。所以严平稳时间序列通常只具有理论意义，在实践中用得更多的是条件比较宽松的宽平稳时间序列。

二、宽平稳（week stationary）

宽平稳是使用序列的特征统计量来定义的一种平稳性。它认为序列的统计性质主要由它的低阶矩决定，所以只要保证序列低阶矩平稳（二阶），就能保证序

列的主要性质近似稳定。

定义 2.2 如果 $\{X_t\}$ 满足如下三个条件：

(1) 任取 $t\in T$，有 $EX_t^2<\infty$；

(2) 任取 $t\in T$，有 $EX_t=\mu$，μ 为常数；

(3) 任取 t，s，$k\in T$，且 $k+s-t\in T$，有 $\gamma(t,s)=\gamma(k,k+s-t)$；

则称 $\{X_t\}$ 为宽平稳时间序列。宽平稳也称为弱平稳或二阶平稳（second-order stationary）。

显然，严平稳比宽平稳的条件严格。严平稳是对序列联合分布的要求，以保证序列所有的统计特征都相同；而宽平稳只要求序列二阶平稳，对于高于二阶的定没有任何要求。所以通常情况下，严平稳序列也满足宽平稳条件，而宽平稳序列不能反推严平稳成立。

但这不是绝对的，两种情况都有特例。

比如服从柯西分布的严平稳序列就不是宽平稳序列，因为它不存在一、二阶矩，所以无法验证它二阶平稳。严格地讲，只有存在二阶矩的严平稳序列才能保证它一定也是宽平稳序列。

宽平稳一般推不出严平稳，但当序列服从多元正态分布时，则二阶平稳可以推出严平稳。

定义 2.3 时间序列 $\{X_t\}$ 称为正态时间序列，如果任取正整数 n，任取 t_1，t_2，…，$t_n\in T$，相对应的有限维随机变量 X_1，X_2，…，X_n 服从 n 维正态分布，密度函数为：

$$f_{t_1,t_2,\cdots,t_n}(\tilde{X}_n)=(2\pi)^{-\frac{n}{2}}\mid \boldsymbol{\Gamma}_n\mid^{-\frac{1}{2}}\exp\left[-\frac{1}{2}(\tilde{X}_n-\tilde{\mu}_n)'\Gamma_n^{-1}(\tilde{X}_n-\tilde{\mu}_n)\right]$$

其中，$\tilde{X}_n=(X_1, X_2, \cdots, X_n)'$，$\tilde{\mu}_n=(EX_1, EX_2, \cdots, EX_n)'$，$\boldsymbol{\Gamma}_n$ 为协方差阵：

$$\boldsymbol{\Gamma}_n=\begin{bmatrix}\gamma(t_1,t_1) & \gamma(t_1,t_2) & \cdots & \gamma(t_1,t_n)\\ \gamma(t_2,t_1) & \gamma(t_2,t_2) & \cdots & \gamma(t_2,t_n)\\ \vdots & \vdots & & \vdots\\ \gamma(t_n,t_1) & \gamma(t_n,t_2) & \cdots & \gamma(t_n,t_n)\end{bmatrix}$$

从正态随机序列的密度函数可以看出，它的 n 维分布仅由均值向量和协方差阵决定，换言之，对正态随机序列而言，只要二阶矩平稳，就等于分布平稳了。所以宽平稳正态时间序列一定是严平稳时间序列。对于非正态过程，就没有这个性质了。

在实际应用中，研究最多的是宽平稳随机序列，以后碰见平稳随机序列，如

果不加特殊注明，指的都是宽平稳随机序列。如果序列不满足平稳条件，就称为非平稳序列。

2.1.3 平稳时间序列的统计性质

根据平稳时间序列的定义，可以推断出它一定具有如下两个重要的统计性质。

一、常数均值

$EX_t=\mu,\ \forall t\in T$

二、自协方差函数和自相关函数只依赖于时间的平移长度而与时间的起止点无关

$$\gamma(t,s)=\gamma(k,k+s-t),\forall t,s,k\in T$$

根据这个性质，可以将自协方差函数 $\gamma(t,s)$ 由二维简化为一维 $\gamma(s-t)$：

$$\gamma(s-t)\hat{=}\gamma(t,s),\forall t,s\in T$$

由此引出延迟 k 自协方差函数的概念。

定义 2.4 对于平稳时间序列 $\{X_t,\ t\in T\}$，任取 t，$t+k\in T$，定义 $\gamma(k)$ 为时间序列 $\{X_t\}$ 的延迟 k 自协方差函数：

$$\gamma(k)=\gamma(t,t+k)$$

根据平稳序列的这个性质，容易推断出平稳随机序列一定具有常数方差：

$$DX_t=\gamma(t,t)=\gamma(0),\forall t\in T$$

由延迟 k 自协方差函数的概念可以等价得到延迟 k 自相关系数的概念：

$$\rho_k=\frac{\gamma(t,t+k)}{\sqrt{DX_t\cdot DX_{t+k}}}=\frac{\gamma(k)}{\sigma_X^2}=\frac{\gamma(k)}{\gamma(0)}$$

容易验证和相关系数一样，自相关系数具有如下三个性质：

(1) 规范性。

$\rho_0=1$ 且 $|\rho_k|\leqslant1,\ \forall k$

(2) 对称性。

$\rho_k=\rho_{-k}$

(3) 非负定性。

对任意正整数 m，相关阵 $\boldsymbol{\Gamma}_m$ 为对称非负定阵。

$$\boldsymbol{\Gamma}_m=\begin{pmatrix}\rho_0 & \rho_1 & \cdots & \rho_{m-1}\\ \rho_1 & \rho_0 & \cdots & \rho_{m-2}\\ \vdots & \vdots & & \vdots\\ \rho_{m-1} & \rho_{m-2} & \cdots & \rho_0\end{pmatrix}$$

值得注意的是 ρ_k 除了具有这三个性质外，它还具有一个特别的性质：非惟一性。

一个平稳时间序列一定惟一决定了它的自相关函数，但一个自相关函数未必惟一对应着一个平稳时间序列。我们在后面的章节中将证明这一点。这个性质就给我们根据样本的自相关系数的特点来确定模型增加了一定的难度。

2.1.4 平稳时间序列的意义

时间序列分析方法作为数理统计学的一个专业分支，它遵循数理统计学的基本原理，都是利用样本信息来推测总体信息。

传统的统计分析通常都拥有如下数据结构，见表 2—1。

表 2—1

随机变量 / 样本	X_1	…	X_m
1	x_{11}	…	x_{m1}
2	x_{12}	…	x_{m2}
⋮	⋮	…	⋮
n	x_{1n}	…	x_{mn}

根据数理统计学常识，显然要分析的随机变量越少越好（m 越小越好），而每个变量获得的样本信息越多越好（n 越大越好）。因为随机变量越少，分析的过程就会越简单，而样本容量越大，分析的结果就会越可靠。

但是时间序列分析的数据结构有它的特殊性。对随机序列 $\{\cdots, X_1, X_2, \cdots, X_t, \cdots\}$ 而言，它在任意时刻 t 的序列值 X_t 都是一个随机变量，而且由于时间的不可重复性，该变量在任意一个时刻只能获得惟一的一个样本观察值。因而时间序列分析的数据结构如下，见表 2—2。

表 2—2

随机变量 / 样本	…	X_1	…	X_t	…
1	…	x_1	…	x_t	…

由于样本信息太少，如果没有其他的辅助信息，通常这种统计结构是没有办法进行分析的。而序列平稳性概念的提出可以有效地解决这个困难。

在平稳序列场合，序列的均值等于常数意味着原本含有可列多个随机变量的均值序列

$$\{\mu_t, t \in T\}$$

变成了只含有一个变量的常数序列

$$\{\mu, t \in T\}$$

原本每个随机变量的均值 μ_t，$t \in T$ 只能依靠惟一的一个样本观察值 x_t 去估计

$$\hat{\mu}_t = x_t$$

现在由于 $\mu_t = \mu$，$\forall t \in T$，于是每一个样本观察值 x_t，$\forall t \in T$，都变成了常数均依 μ 的样本观察值

$$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

这极大地减少了随机变量的个数，并增加了待估变量的样本容量。换句话说，这极大地简化了时序分析的难度，同时也提高了对均值函数的估计精度。

同理，根据平稳序列二阶矩平稳的性质，我们可以得到基于全体观察样本计算出来的延迟 k 自协方差函数的估计值：

$$\hat{\gamma}(k) = \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{n-k}, \forall 0 < k < n$$

并进一步推导出总体方差的估计值：

$$\hat{\gamma}(0) = \frac{\sum_{t=1}^{n} (x_t - \bar{x})^2}{n-k}$$

和延迟 k 自相关系数的估计值：

$$\hat{\rho}_k = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)}, \forall 0 < k < n$$

当延迟阶数 k 远远小于样本容量 n 时：

$$\hat{\rho}_k \cong \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{n} (x_t - \bar{x})^2}, \forall 0 < k \leqslant n$$

2.1.5 平稳性的检验

对序列的平稳性有两种检验方法，一种是根据时序图和自相关图显示的特征

做出判断的图检验方法；一种是构造检验统计量进行假设检验的方法。

图检验方法是一种操作简便、运用广泛的平稳性判别方法，它的缺点是判别结论带有很强的主观色彩。所以最好能用统计检验方法加以辅助判断。目前最常用的平稳性统计检验方法是单位根检验（unit root test）。由于目前知识的局限性，本章将主要介绍平稳性的图检验方法，单位根检验将在第 6 章详细介绍。

一、时序图检验

所谓时序图就是一个平面二维坐标图，通常横轴表示时间，纵轴表示序列取值。时序图可以直观地帮助我们掌握时间序列的一些基本分布特征。

根据平稳时间序列均值、方差为常数的性质，平稳序列的时序图应该显示出该序列始终在一个常数值附近随机波动，而且波动的范围有界的特点。如果观察序列的时序图显示出该序列有明显的趋势性或周期性，那它通常不是平稳序列。根据这个性质，很多非平稳序列通过查看它的时序图可以立刻被识别出来。

【例 2.1】 绘制 1964—1999 年中国纱年产量序列时序图（数据见附录 1.2）。如图 2—1 所示。

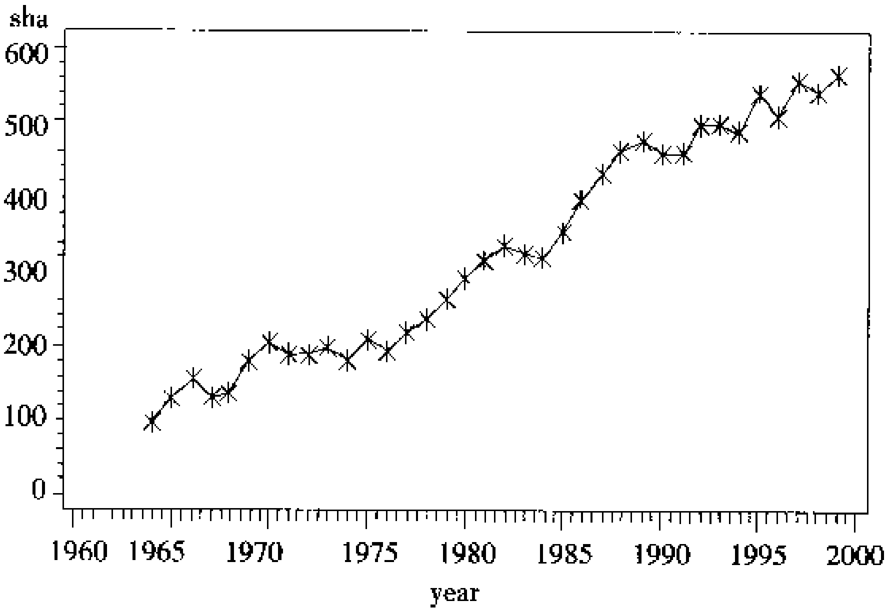


图 2—1 中国纱年产量时序图

时序图给我们提供的信息非常明确，中国纱年产量序列有明显的递增趋势，所以它一定不是平稳序列。

【例 2.2】 绘制 1962 年 1 月至 1975 年 12 月平均每头奶牛月产奶量序列时序

图（数据见附录 1.3)。如图 2—2 所示。

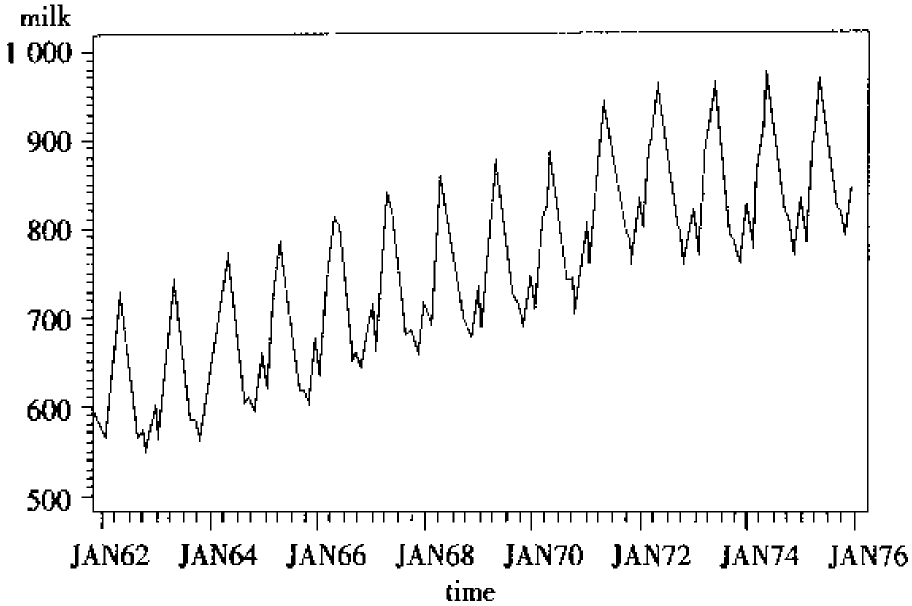


图 2—2 平均每头奶牛月产奶量序列时序图

时序图清晰地显示平均每头牛的月产奶量以年为周期呈现出规则的周期性，除此之外，还有明显的逐年递增的趋势。显然该序列也一定不是平稳序列。

【例 2.3】 绘制 1949—1998 年北京市每年最高气温序列时序图（数据见附录 1.4)。如图 2—3 所示。

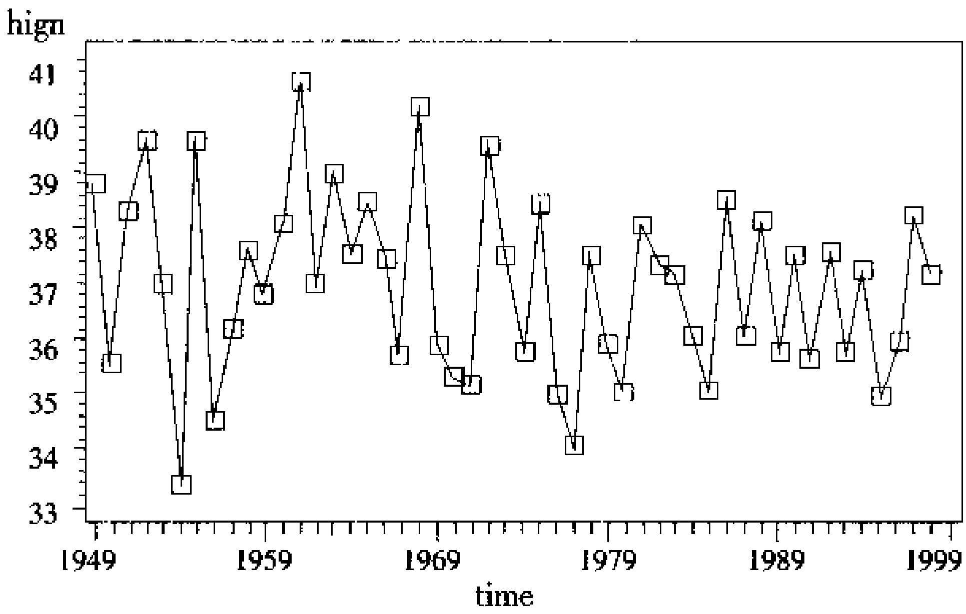


图 2—3 北京市每年的最高温度时序图

时序图显示北京市每年的最高温度始终围绕在 37℃附近随机波动，没有明显趋势或周期，基本可以视为平稳序列。为了稳妥起见，我们还需要利用自相关图进一步辅助识别。

二、自相关图检验

自相关图是一个平面二维坐标悬垂线图，一个坐标轴表示延迟时期数，另一个坐标轴表示自相关系数，通常以悬垂线表示自相关系数的大小。

在后面的章节里我们会证明平稳序列通常具有短期相关性。该性质用自相关系数来描述就是随着延迟期数 k 的增加，平稳序列的自相关系数 $\hat{\rho}_k$ 会很快地衰减向零。反之，非平稳序列的自相关系数 $\hat{\rho}_k$ 衰减向零的速度通常比较慢，这就是我们利用自相关图进行平稳性判断的标准。

【例 2.1 续】 绘制 1964—1999 年中国纱年产量序列自相关图。见图 2—4。

Autocorrelations

Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
0	21 741.103	1.000 00	********************
1	19 869.670	0.913 92	******************
2	18 336.945	0.843 42	*****************
3	16 679.644	0.767 19	***************
4	15 119.827	0.695 45	**************
5	13 234.768	0.608 74	************
6	11 822.365	0.543 78	***********
7	10 355.425	0.476 31	**********
8	8 597.171	0.395 43	********
9	6 977.227	0.320 92	******
10	5 262.589	0.242 06	*****
11	3 185.458	0.146 52	***
12	1 257.065	0.057 82	*
13	−717.129	−.032 98	*
14	−2 356.762	−.108 40	**
15	−3 657.864	−.168 25	***
16	−4 675.021	−.215 03	****
17	−5 645.938	−.259 69	*****
18	−6 662.959	−.306 47	******
19	−7 523.279	−346 04	*******
20	−8 300.856	−.381 80	********
21	−9 068.912	−.417 13	********
22	−9 409.375	−4.32 79	*********

"."marks two standard errors

图 2—4 中国纱年产量序列自相关图

该图横轴表示自相关系数，纵轴表示延迟时期数，用水平方向的垂线表示自相关系数的大小。从图中我们发现序列的自相关系数递减到零的速度相当缓慢，在很长的延迟时期里，自相关系数一直为正，而后，又一直为负，在自相关图上显示出明显的三角对称性，这是具有单调趋势的非平稳序列的一

种典型的自相关图形式。这和该序列时序图（图 2—1）显示的显著的单调递增性是一致的。

【例 2.2 续】 绘制 1962 年 1 月至 1975 年 12 月平均每头奶牛的月产奶量序列自相关图。见图 2—5。

Autocorrelations

Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
0	10 383.588	1.000 00	********************
1	9 257.734	0.891 57	******************
2	8 080.289	0.778 18	****************
3	6 440.643	0.620 27	************
4	5 053.314	0.486 66	**********
5	4 445.713	0.428 15	*********
6	3 904.890	0.376 06	********
7	4 306.827	0.414 77	********
8	4 716.761	0.454 25	*********
9	5 833.655	0.561 81	***********
10	7 128.946	0.686 56	**************
11	7 980.333	0.768 55	***************
12	8 773.234	0.844 91	*****************
13	7 735.639	0.744 99	***************
14	6 621.269	0.637 67	*************
15	5 084.621	0.489 68	**********
16	3 775.004	0.363 55	*******
17	3 176.849	0.305 95	******
18	2 646.859	0.254 91	*****
19	2 984.458	0.287 42	******
20	3 328.659	0.320 57	******
21	4 324.928	0.416 52	********
22	5 489.933	0.528 71	***********
23	6 265.032	0.603 36	************
24	6 986.088	0.672 80	*************

"." marks two standard errors

图 2—5　平均每头奶牛的牛奶月产量序列的自相关图

自相关图显示序列自相关系数长期位于零轴的一边，这是具有单调趋势序列的典型特征，同时自相关图呈现出明显的正弦波动规律，这是具有周期变化规律的非平稳序列的典型特征。自相关图显示出来的这两个性质和该序列时序图（图 2—2）显示出的带长期递增趋势的周期性质是非常吻合的。

【例 2.3 续】 绘制 1949—1998 年北京市每年最高气温序列自相关图。见图 2—6。

自相关图显示该序列的自相关系数一直都比较小，始终控制在 2 倍的标准差范围以内，可以认为该序列自始至终都在零轴附近波动，这是随机性非常强的平稳时间序列通常具有的自相关图性质。

```
                                  Autocorrelations
Lag  Covariance    Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
 0   2.569 604     1.000 00   |                    |********************|
 1   -0.449 960    -.175 11   |                .****|       .            |
 2   -0.009 107 8  -.003 54   |                .    |       .            |
 3   0.463 204     0.180 26   |                .    |****   .            |
 4   0.059 232     0.023 05   |                .    |       .            |
 5   -0.421 428    -.164 00   |                . ***|       .            |
 6   0.253 512     0.098 66   |                .    |**     .            |
 7   -0.067 559    -.026 29   |                .   *|       .            |
 8   -0.008 327 4  -.003 24   |                .    |       .            |
 9   -0.057 247    -.022 28   |                .    |       .            |
10   0.148 917     0.057 95   |                .    |*      .            |
11   0.095 461     0.037 15   |                .    |*      .            |
12   -0.267 799    -.104 22   |                .  **|       .            |
13   0.260 969     0.101 56   |                .    |**     .            |
14   0.011 069     0.004 31   |                .    |       .            |
15   -0.069 243    -.026 95   |                .   *|       .            |
                          "."marks two standard errors
```

图 2—6　1949—1998 年北京市最高气温序列的自相关图

2.2　纯随机性检验

拿到一个观察值序列之后，首先是判断它的平稳性。通过平稳性检验，序列可以分为平稳序列和非平稳序列两大类。

对于非平稳序列，由于它不具有二阶矩平稳的性质，所以对它的统计分析要周折一些，通常要进行进一步的检验、变换或处理之后，才能确定适当的拟合模型。

如果序列平稳，情况就简单多了，我们有一套非常成熟的平稳序列建模方法。但是，并不是所有的平稳序列都值得建模。只有那些序列值之间具有密切的相关关系，历史数据对未来的发展有一定影响的序列，才值得我们花时间去挖掘历史数据中的有效信息，用来预测序列未来的发展。

如果序列值彼此之间没有任何相关性，那就意味着该序列是一个没有记忆的序列，过去的行为对将来的发展没有丝毫影响，这种序列我们称之为纯随机序列。从统计分析的角度而言，纯随机序列是没有任何分析价值的序列。

为了确定平稳序列还值不值得继续分析下去，我们需要对平稳序列进行纯随机性检验。

2.2.1 纯随机序列的定义

定义 2.5 如果时间序列 $\{X_t\}$ 满足如下性质：

(1) 任取 $t\in T$，有 $EX_t=\mu$；

(2) 任取 t，$s\in T$，有

$$\gamma(t,s)=\begin{cases}\sigma^2,t=s\\0,t\neq s\end{cases}$$

称序列 $\{X_t\}$ 为纯随机序列，也称为白噪声（white noise）序列，简记为 $X_t\sim WN(\mu,\sigma^2)$。

之所以称之为白噪声序列，是因为人们最初发现白光具有这种特性。容易证明白噪声序列一定是平稳序列，而且是最简单的平稳序列。

【例 2.4】 随机产生 1 000 个服从标准正态分布的白噪声序列观察值，并绘制时序图。见图 2—7。

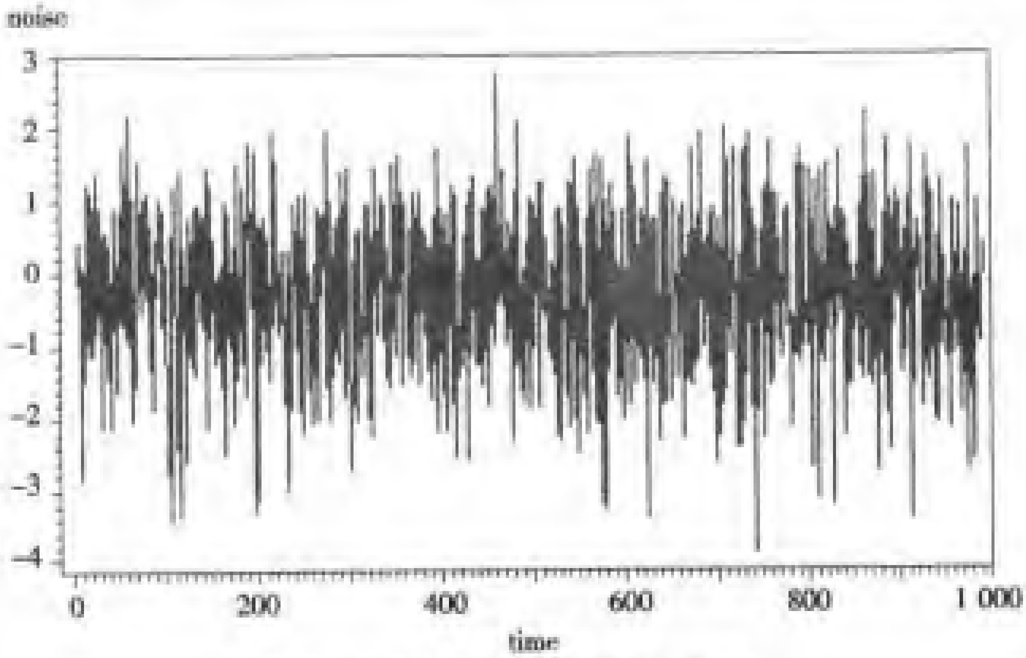


图 2—7 标准正态白噪声序列时序图

2.2.2 白噪声序列的性质

白噪声序列虽然很简单，但它在我们进行时间序列分析时所起的作用却非常大。它的两个重要性质我们在后面的分析过程中要经常用到。

一、纯随机性

由于白噪声序列具有如下性质：

$$\gamma(k)=0,\forall k\neq 0$$

这说明白噪声序列的各项之间没有任何相关关系，这种“没有记忆”的序列就是我们说的纯随机序列。

纯随机序列各项之间没有任何关联，序列在进行完全无序的随机波动。一旦某个随机事件呈现出纯随机运动的特征，我们就认为该随机事件没有包含任何值得提取的有用信息，我们就应该终止分析了。

如果序列值之间呈现出某种显著的相关关系：

$$\gamma(k)\neq 0,\exists k\neq 0$$

就说明该序列不是纯随机序列，该序列间隔 k 期的序列值之间存在着一定程度的相互影响关系，这种相互影响关系，统计上称为相关信息。我们分析的目的就是要想方设法把这种相关信息从观察值序列中提取出来。一旦观察值序列中蕴含的相关信息被我们充分提取出来了，那么剩下的残差序列就应该呈现出纯随机的性质了。所以纯随机性还是我们判断相关信息是否提取充分的一个判别标准。

二、方差齐性

所谓方差齐性，就是指序列中每个变量的方差都相等，即

$$DX_t=\gamma(0)=\sigma^2$$

如果序列不满足方差齐性，我们就称该序列具有异方差性质。

在时间序列分析中，方差齐性是一个非常重要的限制条件。因为根据马尔可夫定理，只有方差齐性假定成立时，我们用最小二乘法得到的未知参数估计值才是准确的、有效的。如果假定不成立，那么最小二乘估计值就不是方差最小线性无偏估计，拟合模型的预测精度会受到很大影响。

所以我们在进行模型拟合时，检验内容之一就是要检验拟合模型的残差是否满足方差齐性假定。如果不满足，那就说明残差序列还不是白噪声序列，即拟合模型没有充分提取随机序列中的相关信息，这时拟合模型的精度是值得怀疑的。在这种场合下，我们通常需要使用适当的条件异方差模型来拟合该序列的发展。

2.2.3 纯随机性检验

纯随机性检验也称为白噪声检验，是专门用来检验序列是否为纯随机序列的一种方法。我们知道如果一个序列是纯随机序列，那它的序列值之间应该没有任何相关关系，即满足

$$\gamma(k)=0,\forall k\neq 0$$

这是一种理论上才会出现的理想状况。实际上，由于观察值序列的有限性，导致纯随机序列的样本自相关系数不会绝对为零。

【例 2.4 续】 绘制例 2.4 标准正态白噪声序列的样本自相关图。如图 2—8 所示。

```
                              Autocorrelations
Lag  Covariance     Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
 0    1.005 528     1.000 00   |                    |********************|
 1   -0.001 064 7   -.001 06   |                   .|.                   |
 2   -0.036 747     -.036 55   |                   *|.                   |
 3   -0.006 256 7   -.006 22   |                   .|.                   |
 4    0.011 938     0.011 87   |                   .|.                   |
 5   -0.025 139     -.025 00   |                   *|.                   |
 6   -0.014 428     -.014 35   |                   .|.                   |
 7    0.008 856 5   0.008 81   |                   .|.                   |
 8   -0.010 179     -.010 12   |                   .|.                   |
 9   -0.026 893     -.026 74   |                   *|.                   |
10   -0.024 882     -.024 75   |                   .|.                   |
11   -0.014 021     -.013 94   |                   .|.                   |
12    0.035 527     0.035 33   |                   .|*                   |
                     "."marks two standard errors
```

图 2—8 白噪声序列样本自相关图

样本自相关图显示这个纯随机序列没有一个样本自相关系数严格等于零。但这些自相关系数确实都非常小，都在零值附近以一个很小的幅度做着随机波动。这就提醒我们应该考虑样本自相关系数的分布性质，从统计意义上来判断序列的性质。

Barlett 证明，如果一个时间序列是纯随机的，得到一个观察期数为 n 的观察序列 $\{x_t,\ t=1,\ 2,\ \cdots,\ n\}$，那么该序列的延迟非零期的样本自相关系数将近似服从均值为零、方差为序列观察期数倒数的正态分布，即

$$\hat{\rho}_k \dot{\sim} N(0,\frac{1}{n}),\forall k\neq 0$$

式中，n 为序列观察期数。

根据 Barlett 定理，我们可以构造检验统计量来检验序列的纯随机性。

一、假设条件

由于序列值之间的变异性是绝对的，而相关性是偶然的，所以假设条件如下确定。

原假设：延迟期数小于或等于 m 期的序列值之间相互独立。

备择假设：延迟期数小于或等于 m 期的序列值之间有相关性。

该假设条件用数学语言描述即为：

$$H_0:\rho_1=\rho_2=\cdots=\rho_m=0,\forall m\geqslant 1$$

$$H_1:\text{至少存在某个 }\rho_k\neq 0,\forall m\geqslant 1,k\leqslant m$$

二、检验统计量

1. Q统计量

为了检验这个联合假设，Box 和 Pierce 推导出了Q统计量：

$$Q = n\sum_{k=1}^{m}\hat{\rho}_k^2$$

式中，n 为序列观测期数；m 为指定延迟期数。

根据正态分布和卡方分布之间的关系，我们很容易推导出Q统计量近似服从自由度为m 的卡方分布：

$$Q = n\sum_{k=1}^{m}\hat{\rho}_k^2 \sim \chi^2(m)$$

当Q统计量大于$\chi^2_{1-\alpha}(m)$分位点，或该统计量的P 值小于α 时，则可以以 $1-\alpha$ 的置信水平拒绝原假设，认为该序列为非白噪声序列；否则，接受原假设，认为该序列为纯随机序列。

2. LB统计量

在实际应用中人们发现Q统计量在大样本场合（n 很大的场合）检验效果很好，但在小样本场合就不太精确。为了弥补这一缺陷，Box 和 Ljung 又推导出 LB（Ljung-Box）统计量：

$$LB = n(n+2)\sum_{k=1}^{m}\left(\frac{\hat{\rho}_k^2}{n-k}\right)$$

式中，n 为序列观测期数；m 为指定延迟期数。Box 和 Ljung 证明 LB 统计量同样近似服从自由度为m 的卡方分布。

实际上LB统计量就是Box 和 Pierce 的Q统计量的修正，所以人们习惯把它们统称为Q统计量，分别记作Q_{BP}统计量（Box 和 Pierce 的Q统计量）和Q_{LB}统计量（Box 和 Ljung 的Q统计量），在各种检验场合普遍采用的Q统计量通常指的都是LB统计量。

【例 2.4 续】 计算例 2.4 中白噪声序列的延迟 6 期、延迟 12 期的Q_{LB}统计量的值，并判断该序列的随机性（α=0.05）。

由图 2--8 我们可以得到该序列延迟 12 期样本自相关系数，数据如下，见表 2--3。

表 2—3

延迟期数 k	1	2	3	4	5	6
$\hat{\rho}_k$	−0.001	−0.037	−0.006	0.012	−0.025	−0.014
延迟期数 k	7	8	9	10	11	12
$\hat{\rho}_k$	0.009	−0.010	−0.027	−0.025	−0.014	0.035

根据如上数据，很容易计算出表 2—4 的结果。

表 2—4

延迟	Q_{LB}统计量检验	
	Q_{LB}统计量值	P 值
延迟 6 期	2.36	0.883 8
延迟 12 期	5.35	0.945 4

由于 P 值显著大于显著性水平 α，所以该序列不能拒绝纯随机的原假设。换言之，我们可以认为该序列的波动没有任何统计规律可循。因而，可以停止对该序列的统计分析。

还需要解释的一点是，为什么在本例只检验了前 6 期和前 12 期延迟的 Q 统计量和 LB 统计量就直接判断该序列是白噪声序列呢？为什么不进行全部 999 期延迟检验呢？

这是因为平稳序列通常具有短期相关性，如果序列值之间存在显著的相关关系，通常只存在在延迟时期比较短的序列值之间。所以，如果一个平稳序列短期延迟的序列值之间都不存在显著的相关关系，通常长期延迟之间就更不会存在显著的相关关系。

另一方面，假如一个平稳序列显示出显著的短期相关性，那么该序列就一定不是白噪声序列，我们就可以对序列值之间存在的相关性进行分析。假如此时考虑的延迟时期数太长，反而可能淹没了该序列的短期相关性。因为平稳序列只要延迟时期足够长，自相关系数都会收敛于零。

【例 2.3 续】 对 1949—1998 年北京市最高气温序列做白噪声检验。(α=0.05)

检验结果见表 2—5。

表 2—5

延迟	LB 统计量检验	
	LB 统计量值	P 值
延迟 6 期	5.58	0.471 3
延迟 12 期	6.71	0.876 0

根据这个检验结果，不能拒绝序列纯随机的原假设。因而可以认为北京市最高气温的变动属于纯随机波动。这说明我们很难根据历史信息预测未来年份的最高气温。至此，对该序列的分析也就结束了。

【例 2.5】 对 1950—1998 年北京市城乡居民定期储蓄所占比例序列的平稳性与纯随机性进行检验（数据见附录 1.5）。

1. 绘制该序列时序图

见图 2—9。

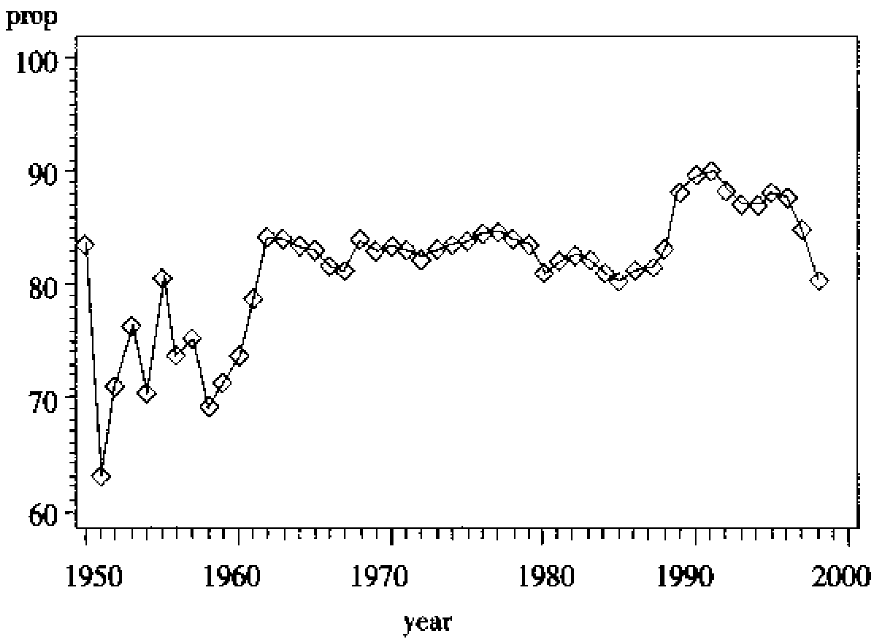


图 2—9 北京市城乡居民定期储蓄序列时序图

该时序图显示北京市城乡居民的定期储蓄始终占储蓄存款余额的 80%左右，波动比较平稳。

2. 自相关图检验

考察该序列的样本自相关图，进一步检验该序列的平稳性。见图 2—10。

```
                              Autocorrelations

Lag  Covariance   Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  0  30.725 523   1.000 00   |                    |********************|
  1  21.583 411   0.702 46   |              .     |**************      |
  2  18.293 557   0.595 39   |             .      |************        |
  3  14.684 303   0.477 92   |            .       |**********          |
  4  10.080 193   0.328 07   |           .        |*******    .        |
  5  10.931 717   0.355 79   |           .        |*******    .        |
  6   9.318 240   0.303 27   |          .         |******      .       |
  7   8.944 975   0.291 13   |          .         |******      .       |
  8   4.927 541   0.160 37   |          .         |***         .       |
  9   1.842 114   0.059 95   |          .         |*           .       |
 10  -1.151 434   -.037 47   |          .        *|            .       |
 11  -2.369 343   -.077 11   |          .       **|            .       |
 12  -1.130 247   -.036 79   |          .        *|            .       |

                     "."marks two standard errors
```

图 2—10 北京市城乡居民定期储蓄序列自相关图

样本自相关图显示延迟 3 阶之后，自相关系数都落入 2 倍标准差范围以内，而且自相关系数向零衰减的速度非常快，延迟 8 阶之后自相关系数即在零值附近波动。这是一个非常典型的短期相关的样本自相关图。由时序图和样本自相关图的性质，可以认为该序列平稳。

3. 纯随机性检验（α=0.05）

检验结果见表 2- 6。

表 2—6

延迟阶数	LB统计量检验	
	LB检验统计量的值	P值
6	75.46	<0.000 1
12	82.57	<0.000 1

检验结果显示，在各阶延迟下LB检验统计量的 P 值都非常小（<0.000 1），所以我们可以以很大的把握（置信水平>99.999%）断定北京市城乡居民定期储蓄比例序列属于非白噪声序列。

结合前面的平稳性检验结果，说明该序列不仅可以视为是平稳的，而且还蕴含着值得我们提取的相关信息。这种平稳非白噪声序列是目前最容易分析的一种序列，下一章我们就要详细介绍对这种平稳非白噪声序列的建模及预测方法。

2.3 习　题

1. 考虑系列｛1，2，3，4，5，…，20｝：

(1) 判断该系列是否平稳。

(2) 计算该序列的样本自相关系数，$\hat{\rho}_k$（k=1，2，…，6）。

(3) 绘制该样本自相关图，并解释该图形。

2. 1975—1980 年夏威夷岛莫那罗亚火山（Mauna Loa）每月释放的 CO_2 数据如下（单位：ppm），见表 2 -7。

表 2—7

330.45	330.97	331.64	332.87	333.61	333.55
331.90	330.05	328.58	328.31	329.41	330.63
331.63	332.46	333.36	334.45	334.82	334.32
333.05	330.87	329.24	328.87	330.18	331.50
332.81	333.23	334.55	335.82	336.44	335.99
334.65	332.41	331.32	330.73	332.05	333.53
334.66	335.07	336.33	337.39	337.65	337.57
336.25	334.39	332.44	332.25	333.59	334.76
335.89	336.44	337.63	338.54	339.06	338.95
337.41	335.71	333.68	333.69	335.05	336.53
337.81	338.16	339.88	340.57	341.19	340.87
339.25	337.19	335.49	336.63	337.74	338.36

(1) 绘制该序列时序图，并判断该系列是否平稳。

(2) 计算该序列的样本自相关系数 $\hat{\rho}_k$（$k=1, 2, \cdots, 24$）。

(3) 绘制该样本自相关图，并解释该图形。

3. 1945—1950 年费城月度降雨量数据如下（单位：mm），见表 2—8。

表 2—8

69.3	80.0	40.9	74.9	84.6	101.1	225.0	95.3	100.6	48.3	144.5	128.3
38.4	52.3	68.6	37.1	148.6	218.7	131.6	112.8	81.8	31.0	47.5	70.1
96.8	61.5	55.6	171.7	220.5	119.4	63.2	181.6	73.9	64.8	166.9	48.0
137.7	80.5	105.2	89.9	174.8	124.0	86.4	136.9	31.5	35.3	112.3	143.0
160.8	97.0	80.5	62.5	158.2	7.6	165.9	106.7	92.2	63.2	26.2	77.0
52.3	105.4	144.3	49.5	116.1	54.1	148.6	159.3	85.3	67.3	112.8	59.4

(1) 计算该序列的样本自相关系数 $\hat{\rho}_k$（$k=1, 2, \cdots, 24$）。

(2) 判断该序列的平稳性。

(3) 判断该序列的纯随机性。

4. 若序列长度为 100，前 12 个样本自相关系数如下：

$\rho_1=0.02$　$\rho_2=0.05$　$\rho_2=0.10$　$\rho_4=-0.02$　$\rho_5=0.05$　$\rho_6=0.01$

$\rho_7=0.12$　$\rho_8=-0.06$　$\rho_6=0.08$　$\rho_{10}=-0.05$　$\rho_{11}=0.02$　$\rho_{12}=-0.05$

该序列能否视为纯随机序列？

5. 表 2—9 数据是某公司在 2000—2003 年期间每月的销售量。

表 2—9

月份	2000 年	2001 年	2002 年	2003 年
1	153	134	145	117
2	187	175	203	178
3	234	243	189	149
4	212	227	214	178
5	300	298	295	248
6	221	256	220	202
7	201	237	231	162
8	175	165	174	135
9	123	124	119	120
10	104	106	85	96
11	85	87	67	90
12	78	74	75	63

(1) 绘制该序列时序图及样本自相关图。

(2) 判断该序列的平稳性。

(3) 判断该序列的纯随机性。

6. 1969 年 1 月至 1973 年 9 月在芝加哥海德公园内每 28 天发生的抢包案件数见表 2—10。

表 2—10

10	15	10	10	12	10	7	7	10	14	8	17
14	18	3	9	11	10	6	12	14	10	25	29
33	33	12	19	16	19	19	12	34	15	36	29
26	21	17	19	13	20	24	12	6	14	6	12
9	11	17	12	8	14	14	12	5	8	10	3
16	8	8	7	12	6	10	8	10	5		

(1) 判断该序列 $\{x_t\}$ 的平稳性及纯随机性。

(2) 对该序列进行函数运算：

$$y_t = x_t - x_{t-1}$$

并判断序列 $\{y_t\}$ 的平稳性及纯随机性。

2.4 上机指导

2.4.1 绘制时序图

在 SAS 系统中，使用 GPLOT 程序可以绘制多种精美的时序图，以表 2—11 数据为例，介绍 GPLOT 程序的基本命令。

表 2—11

Time	Price1	Price2
2004 年 7 月	12.85	15.21
2004 年 8 月	13.29	14.23
2004 年 9 月	12.41	14.69
2004 年 10 月	15.21	13.27
2004 年 11 月	14.23	16.75
2004 年 12 月	13.56	15.33

```
data example2_1;
input price1 price2;
time=intnx ('month','01jul2004'd, _n_-1);
```

```
format time date. ;
cards;
12.85   15.21
13.29   14.23
12.41   14.69
15.21   13.27
14.23   16.75
13.56   15.33
;
proc gplot data=example2_1;
plot price1 * time=1 price2 * Time=2/overlay;
symbol1 c=black v=star i=join;
symbol2 c=red v=circle i=needle;
run;
```

语句说明：

(1) "proc gplot data=example2_1;"

这是告诉系统，下面将准备对临时数据集 example2_1 中的数据绘图。

(2) "plot price1 * time=1 price2 * time=2/overlay;"

这也是要求系统要绘制两条时序曲线，第一条是以 price1 为纵坐标，time 为横坐标，以 symbol1 语句所规定的格式绘制。第二条是以 price2 为纵坐标，time 为横坐标，以 symbol2 语句所规定的格式绘制。overlay 选项指令系统将这两条时序线绘制在同一张图中，同时显示。如果没有 overlay 选项，系统将这两条时序线分开输出。

(3) "symbol1 c=black v=star i=join;"

symbol 语句是专门指令绘制的格式，一个 GPLOT 程序中允许使用多个 symbol 语句，所以就有了 symbol1，symbol2，…。

symbol 语句中有许多选项，最常用的三大选项是：

C——图线颜色，可自由选择 red（红色），black（黑色），green（绿色）blue（蓝色），pink（粉红色）等各种颜色。

V——表示观察值的图形，可自由选择 star（星号），dot（点），circle（圆圈），diamond（菱形）等各种形状，也可选择 none（不使用特别图形标注观察值）。

I——观察值之间的连线方式，可自由选择 join（线性连接），spline（光滑

连接），needle（作观察值到横轴的悬垂线）等各种连线方式，也可选择 none（不作任何连接）。

本例输出的时序图见图 2—11。

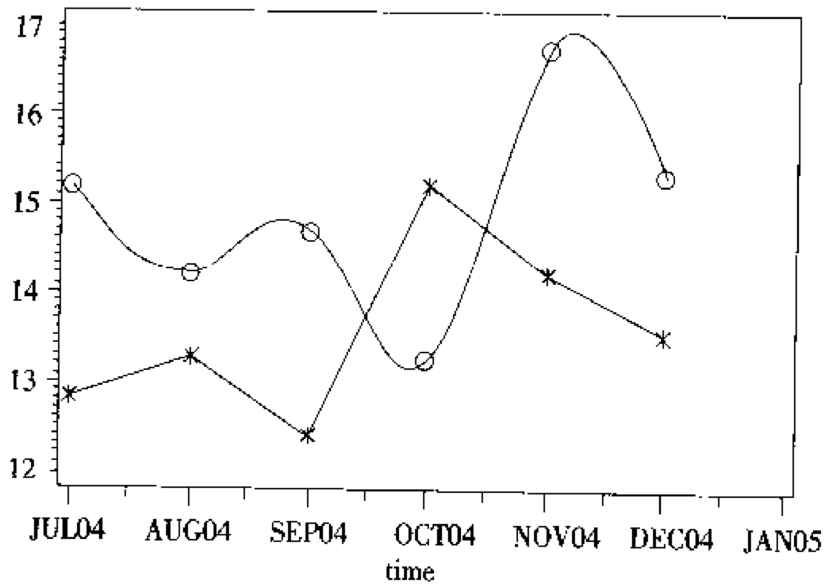


图 2—11 两时间序列重叠显示时序图

2.4.2 平稳性与纯随机性检验

一、平稳性检验

为了判断序列是否平稳，除了需要考察时序图的性质，还需要对自相关图进行检验。SAS 系统 ARIMA 过程中的 IDENTIFY 语句可以提供非常醒目的自相关图。

以临时数据集 example2 _ 2 中的数据为例，介绍 IDENTIFY 语句的使用。

```
data example2 _ 2;
input freq@@;
year=intnx ('year','1jan1970'd, _ n _-1);
format year year4. ;
cards;
97 154 137.7 149 164 157 188 204 179 210 202 218 209
204 211 206 214 217 210 217 219 211 233 316 221 239
215 228 219 239 224 234 227 298 332 245 357 301 389
;
proc arima data=example2 _ 2;
```

identify var=freq;

run;

语句说明：

(1) “proc arima data=example2 _ 2;”

这条命令是告诉系统，下面要对临时数据集 example2 _ 2 中的数据进行 ARIMA 程序分析。

(2) “identify var=freq;”

这是指令系统对变量 freq 的某些重要性质进行识别。每条 IDENTIFY 语句执行之后，会给出五方面的信息：

1) 分析变量的描述性统计；

2) 样本自相关图；

3) 样本偏自相关图；

4) 样本逆自相关图；

5) 纯随机检验结果。

执行本例程序，IDENTIFY 语句输出的描述性信息如下：

```
        Name of Variable=freq
Mean of Woking Series         222.941
Standard Deviation           56.82834
Number of Observations             39
```

这部分给出了分析变量的名称、序列均值、标准差和观察值个数。

IDENTIFY 语句输出结果的第二部分为自相关图（autocorrelations），本例获得的样本自相关图见图 2—12。

```
                                      Autocorrelations
Lag Covariance Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1  Std Error
  0  3 229.460    1.000 00  |                    |********************|         0
  1  1 865.267    0.577 58  |              .     |************        |  0.160 128
  2  1 933.397    0.598 67  |            .       |************        |  0.206 757
  3  1 299.536    0.402 40  |          .         |******** .          |  0.247 242
  4  1 238.607    0.383 53  |         .          |********  .         |  0.263 501
  5    877.946    0.271 86  |         .          |*****     .         |  0.277 446
  6    439.920    0.136 22  |         .          |***       .         |  0.284 194
  7    375.215    0.116 19  |         .          |**        .         |  0.285 863
  8    331.853    0.102 76  |         .          |**        .         |  0.287 071
  9    292.384    0.090 54  |        .           |**         .        |  0.288 013

                               "."marks two standard errors
```

图 2—12 序列 FREQ 样本自相关图

其中：

Lag——延迟阶数。

Covariance——延迟阶数给定后的自协方差函数。

Correlation——延迟阶数给定后的自相关函数。

STD ERROR——自相关函数的标准差。

“.”——2 倍标准差范围。

二、纯随机性检验

为了判断序列是否有分析价值，我们必须对序列进行纯随机性检验，即白噪声检验。ARIMA 过程中的 IDENTIFY 语句同时提供了这方面的信息。

在 IDENTIFY 输出结果的最后一部分信息就是白噪声检验结果。本例中白噪声检验输出结果如下

Autocorrelation Check for White Noise

To Lag	Chi-Square	DF	Pr> ChiSq	······	······	Autocorrelations	······	······	······
6	47.81	6	<.000 1	0.578	0.599	0.402	0.384	0.272	0.136

其中：

To Lag——是延迟阶数。

Chi-Square——是 Q_{LB} 统计量的值，该统计量服从 χ^2 分布。

Df——是 Q_{LB} 统计量服从的 χ^2 分布的自由度。

Pr>Chisq——该 Q_{LB} 统计量的 P 值。

Autocorrelation——计算得出的延迟各阶 Q_{LB} 统计量的样本自相关系数的具体数值。

第3章

平稳时间序列分析

一个序列经过预处理被识别为平稳非白噪声序列，那就说明该序列是一个蕴含着相关信息的平稳序列。在统计上，我们通常是建立一个线性模型来拟合该序列的发展，借此提取该序列中的有用信息。ARMA（auto regression moving average）模型是目前最常用的平稳序列拟合模型。

3.1 方法性工具

在时间序列分析中有一些方法性工具经常被使用，它们可以使我们的模型表达和序列分析更加简洁、方便。所以在介绍具体的模型之前我们先简单介绍这些常用的方法性工具。

3.1.1 差分运算

一、p 阶差分

相距一期的两个序列使之间的减法运算称为1阶差分运算。记∇x_t 为 x_t 的1阶差分：

$$\nabla x_t = x_t - x_{t-1}$$

对 1 阶差分后序列再进行一次 1 阶差分运算称为 2 阶差分。记$\nabla^2 x_t$ 为 x_t 的 2 阶差分：

$$\nabla^2 x_t = \nabla x_t - \nabla x_{t-1}$$

依此类推，对 $p-1$ 阶差分后序列再进行一次 1 阶差分运算称为 p 阶差分。记$\nabla^p x_t$ 为 x_t 的 p 阶差分：

$$\nabla^p x_t = \nabla^{p-1} x_t - \nabla^{p-1} x_{t-1}$$

二、k 步差分

相距 k 期的两个序列值之间的减法运算称为 k 步差分运算。记$\nabla_k x_t$ 为 x_k 的 k 步差分：

$$\nabla_k = x_t - x_{t-k}$$

3.1.2 延迟算子

一、定义

延迟算子类似于一个时间指针，当前序列值乘以一个延迟算子，就相当于把当前序列值的时间向过去拨了一个时刻。记 B 为延迟算子，有

$$x_{t-1} = Bx_1$$

$$x_{t-2} = B^2 x_t$$

$$\vdots$$

$$x_{t-p} = B^p x_t$$

延迟算子有如下性质：

1. $B^0 = 1$
2. 若 c 为任一常数，有 $B(c \cdot x_t) = c \cdot B(x_t) = c \cdot x_{t-1}$
3. 对任意两个序列$\{x_t\}$和$\{y_t\}$，有 $B(x_t \pm y_t) = x_{t-1} \pm y_{t-1}$
4. $B^n x_t = x_{t-n}$
5. $(1-B)^n = \sum_{i=0}^{n} (-1)^n C_n^i B^i$，其中 $C_n^i = \frac{n!}{i!(n-i)!}$

二、用延迟算子表示差分运算

1. p 阶差分

$$\nabla^p x_t = (1-B)^p x_t = \sum_{i=0}^{p} (-1)^p C_p^i x_{t-i}$$

2. k 步差分

$$\nabla_k = x_t - x_{t-k} = (1-B^k) x_r$$

3.1.3 线性差分方程

一、线性差分方程的定义

定义 3.1 称如下形式的方程为序列$\{z_t, t=0, \pm 1, \pm 2, \cdots\}$的线性差分方程：

$$z_t + a_1 z_{t-1} + a_2 z_{t-2} + \cdots + a_p z_{t-p} = h(t) \tag{3.1}$$

式中，$p \geqslant 1$；$a_1, a_2, \cdots, a_p$ 为实数；$h(t)$为 t 的已知函数。

特别地，若 $h(t)=0$，则差分方程

$$z_t + a_1 z_{t-1} + a_2 z_{t-2} + \cdots + a_p z_{t-p} = 0 \tag{3.2}$$

被称为齐次线性差分方程。否则，线性差分方程（3.1）称为非齐次线性差分方程。

二、齐次线性差分方程的解

齐次线性差分方程的求解要借助它的特征方程和特征根。齐次线性差分方程（3.2）的特征方程为：

$$\lambda^p + a_1 \lambda^{p-1} + a_2 \lambda^{p-2} + \cdots + a_p = 0 \tag{3.3}$$

这是一个一元 p 次线性方程，它应该有至少 p 个非零根，称这 p 个非零根为特征方程的特征根，不妨记作

$$\lambda_1, \lambda_2, \cdots, \lambda_p$$

特征根的取值不同，齐次线性差分方程（3.2）的解会有不同的表达式。下面分情况讨论。

1. $\lambda_1, \lambda_2, \cdots, \lambda_p$ 为 p 个不同的实数

这时齐次线性差分方程（3.2）的解为：

$$z_t = c_1 \lambda_1^t + c_2 \lambda_2^t + \cdots + c_p \lambda_p^t$$

式中，c_1，c_2，…，c_p 为任意实数。

2. $\lambda_1, \lambda_2, \cdots, \lambda_p$ 中有相同实根

不妨假定 $\lambda_1 = \lambda_2 = \cdots = \lambda_d$ 为 d 个相同实根，而 $\lambda_{d+1}, \lambda_{d+2}, \cdots, \lambda_p$ 为互不相同实根。这时齐次线性差分方程（3.2）的解为：

$$z_t = (c_1 + c_2 t + \cdots + c_d t^{d-1}) \lambda_1^t + c_{d+1} \lambda_{d+1}^t + \cdots + c_p \lambda_p^t$$

式中，$c_1, c_2, \cdots, c_p$ 为任意实数。

3. $\lambda_1, \lambda_2, \cdots, \lambda_p$ 中有复根

由于差分方程的系数 $a_1, a_2, \cdots, a_p$ 为实数，所以其复根必呈共轭出现。不妨假定

$$\lambda_1 = a + ib = re^{i\bar{\omega}}，\lambda_2 = a - ib = re^{-i\bar{\omega}}$$

为一对共轭复根，其中 $r = \sqrt{a^2 + b^2}$，$\bar{\omega} = \arccos \dfrac{a}{r}$，而 $\lambda_3, \lambda_4, \cdots, \lambda_p$ 为互不相同实根。这时齐次线性差分方程（3.2）的解为：

$$
\begin{aligned}
z_t &= c_1\lambda_1^t + c_2\lambda_2^t + \cdots + c_p\lambda_p^t \\
&= r^t(c_1 e^{it\bar{\omega}} + c_2 e^{-it\bar{\omega}}) + c_3\lambda_3^t + \cdots + c_p\lambda_p^t
\end{aligned}
$$

式中，$c_1, c_2, \cdots, c_p$ 为任意实数。

三、非齐次线性差分方程的解

求非齐次线性差分方程（3.1）的解通常需要进行两步运算。首先求出齐次线性差分方程（3.2）的通解 z'_t；然后求出该非齐次线性差分方程（3.1）的一个特解 z''_t。

所谓特解就是任意一个值 z''_t，使得非齐次线性差分方程（3.1）成立，即

$$z''_t + a_1 z''_{t-1} + a_2 z''_{t-2} + \cdots + a_p z''_{t-p} = h(t)$$

而非齐次线性差分方程（3.1）的通解即为齐次线性差分方程（3.2）的通解 z'_t 和非齐次线性差分方程的特解 z''_t 之和，即

$$z_t = z'_t + z''_t$$

四、时间序列模型与线性差分方程

线性差分方程在时间序列分析中有着重要的应用。常用的时间序列模型和某些模型的自协方差函数和自相关函数都可以视为线性差分方程，而线性差分方程对应的特征根的性质对判断模型的平稳性有着非常重要的意义。

3.2 *ARMA* 模型的性质

ARMA 模型的全称是自回归移动平均（auto regression moving average）模型，它是目前最常用的拟合平稳序列的模型。它又可以细分为 *AR* 模型（auto regression model）、*MA* 模型（moving average model）和 *ARMA* 模型（auto regression moving average model）三大类。

3.2.1 *AR* 模型

一、定义

定义 3.2 具有如下结构的模型称为 p 阶自回归模型，简记为 $AR(p)$：

$$
\begin{cases}
x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + \varepsilon_t \\
\phi_p \neq 0 \\
E(\varepsilon_t) = 0, \mathrm{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t\varepsilon_s) = 0, s \neq t \\
Ex_s\varepsilon_t = 0, \forall s < t
\end{cases}
\tag{3.4}
$$

$AR(p)$ 模型有三个限制条件：

条件一：$\phi_p \neq 0$，这个限制条件保证了模型的最高阶数为 p。

条件二：$E(\varepsilon_t)=0$，$Var(\varepsilon_t)=\sigma_\varepsilon^2$，$E(\varepsilon_t\varepsilon_s)=0$，$s\neq t$，这个限制条件实际上是要求随机干扰序列$\{\varepsilon_t\}$为零均值白噪声序列。

条件三：$Ex_s\varepsilon_t=0$，$\forall s<t$，这个限制条件说明当期的随机干扰与过去的序列值无关。

通常会缺省默认方程（3.4）的限制条件，把 $AR(p)$模型简记为：

$$x_t=\phi_0+\phi_1 x_{t-1}+\phi_2 x_{t-2}+\cdots+\phi_p x_{t-p}+\varepsilon_t \tag{3.5}$$

当 $\phi_0=0$ 时，自回归模型（3.4）又称为中心化 $AR(p)$模型。非中心化 $AR(p)$序列都可以通过下面的变换转化为中心化 $AR(p)$系列。

令

$$\mu=\frac{\phi_0}{1-\phi_0-\cdots-\phi_p},\ y_t=x_t-\mu$$

则$\{y_t\}$为$\{x_t\}$的中心化序列。中心化变换实际上就是非中心化的序列整个平移了一个常数位移，这种整体移动对序列值之间的相关关系没有任何影响，所以今后在分析 AR 模型的相关关系时，都简化为对它的中心化模型进行分析。

引进延迟算子，中心化 $AR(p)$模型又可以简记为：

$$\Phi(B)x_t=\varepsilon_t \tag{3.6}$$

式中，$\Phi(B)=1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p$，称为 p 阶自回归系数多项式。

二、AR 模型平稳性判别

要拟合一个平稳序列的发展，用来拟合的模型显然也应该是平稳的。AR 模型是常用的平稳序列的拟合模型之一，但并非所有的 AR 模型都是平稳的。

【例 3.1】 考察如下四个 AR 模型的平稳性：

（1）$x_t=0.8x_{t-1}+\varepsilon_t$　　（2）$x_t=-1.1x_{t-1}+\varepsilon_t$

（3）$x_t=x_{t-1}-0.5x_{t-2}+\varepsilon_t$　　（4）$x_t=x_{t-1}+0.5x_{t-1}+\varepsilon_t$

拟合这四个序列的序列值，并绘制时序图，见图 3—1。

显然，（1）、（3）模型平稳，（2）、（4）模型非平稳。图示法只是一种粗糙的直观判别方法，我们有两种准确的平稳性判别方法：特征根判别和平稳域判别。

1. 特征根判别

任一个中心化 $AR(p)$ 模型 $\Phi(B)x_t=\varepsilon_t$ 都可以模为一个非齐次线性差分方程：

$$x_t-\phi_1 x_{t-1}-\phi_2 x_{t-2}-\cdots-\phi_p x_{t-p}=\varepsilon_t \tag{3.7}$$

方程（3.7）的通解为：

$$x_t=x_t'+x_t''$$

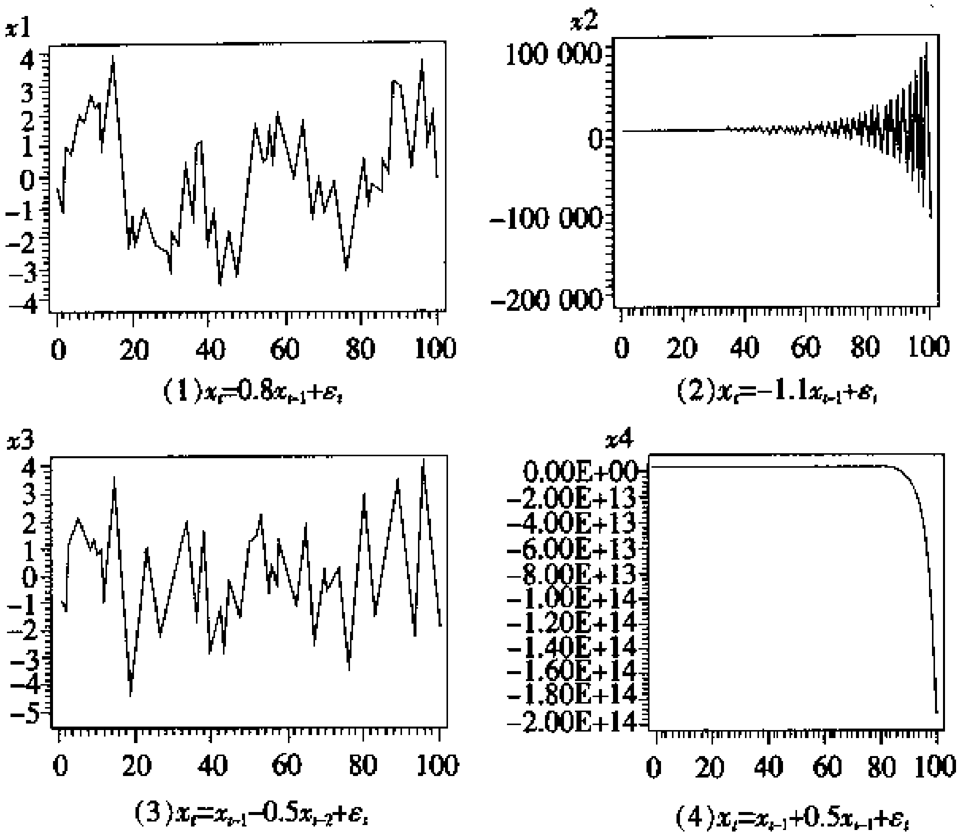


图 3—1　四个 *AR* 序列时序图

式中，x_t'为齐次线性差分方程 $\Phi(B)x_t=0$ 的通解；x_t''为方程（3.7）的一个特解。

（1）求齐次线性差分方程 $\Phi(B)x_t=0$ 的一个通解 x_t'。假定 λ_1，λ_2，…，λ_p 是该特征方程的 p 个特征根。为了有代表性，不妨假设这 p 个特征根取值如下：

$\lambda_1=\lambda_2=\cdots=\lambda_p$ 为 d 个相等实根。

λ_{d+1}，λ_{d+2}，…，λ_{p-2m}为 $p-d-2m$ 个互不等实根。

$\lambda_{j1}=r_je^{i\omega_j}$，$\lambda_{j2}=r_je^{-i\omega_j}$，$j=1$，…，$m$ 为 m 对共轭复根。

那么齐次线性差分方程 $\Phi(B)x_t=0$ 的通解为：

$$x_t'=\sum_{j=1}^{d}c_jt^{j-1}\lambda_1^t+\sum_{j=d+1}^{p-2m}c_j\lambda^t+\sum_{j=1}^{m}\gamma_j^t(c_{1j}\cos t\omega_j+c_{2j}\sin t\omega_j)$$

式中，$c_1,\cdots,c_p,c_{1j},c_{2j}(j=1,\cdots,m)$为任意实数。

（2）求非齐次线性差分方程 $\Phi(B)x_t=\varepsilon_t$ 的一个特解 x_t''。首先，可以证明 $AR(p)$ 模型的自回归系数多项式方程 $\Phi(u)=0$ 的根是齐次线性差分方程 $\Phi(B)x_t=0$ 的特征根的倒数。

证明：设 $\lambda_1,\lambda_2,\cdots,\lambda_p$ 为齐次线性差分方程 $\Phi(B)x_t=0$ 的 p 个特征根，任取 $\lambda_i,i\in(1,2,\cdots,p)$，带入特征方程，有

$$\lambda_i^p-\phi_1\lambda_i^{p-1}-\phi_2\lambda_i^{p-2}-\cdots-\phi_p=0$$

把 $u_i=\frac{1}{\lambda_i}$ 带入 $AR(p)$ 模型的自回归系数多项式，有

$$\Phi(u_i)=1-\phi_1\frac{1}{\lambda_1}-\cdots-\phi_p\frac{1}{\lambda_i^p}$$

$$=\frac{1}{\lambda_i^p}[\lambda_i^p-\phi_1\lambda_i^{p-1}-\cdots-\phi_p]=0$$

根据这个性质，$\Phi(B)$ 可以因子分解成

$$\Phi(B)=\prod_{i=1}^{p}(1-\lambda_i B)$$

由此可以得到非齐次线性差分方程（3.7）的一个特解为：

$$x''_t=\frac{\varepsilon_t}{\Phi(B)}=\frac{\varepsilon_t}{\prod\limits_{i=1}^{p}(1-\lambda_i B)}=\sum_{i=1}^{p}\frac{k_i}{1-\lambda_i B}\varepsilon_t$$

式中，k_i（$i=1$，2，⋯，p）为常数。

（3）求非齐次线性差分方程（3.7）的通解 x_t。

$$x_t=x'_t+x''_t$$

$$=\sum_{j=1}^{p}c_j t^{j-1}\lambda_1^t+\sum_{j=d+1}^{p-2m}c_j\lambda^t+\sum_{j=1}^{m}\gamma_j^t(c_{1j}\cos t\omega_j+c_{2j}\sin t\bar{\omega})+\sum_{i=1}^{p}\frac{k_i}{1-\lambda_i B}\varepsilon_t$$

要使得中心化 $AR(p)$ 模型平稳，即要求对任意实数 $c_1,\cdots,c_{p-2m},c_{1j},c_{2j}(j=1,\cdots,m)$ 有

$$\lim_{t\to\infty}x_t=0 \tag{3.8}$$

式（3.8）成立的充要条件是

$$\begin{aligned}&|\lambda_i|<1,i=1,2,\cdots,p-2m\\&|\gamma_i|<1,i=1,2,\cdots,m\end{aligned} \tag{3.9}$$

式（3.9）实际上就是要求 $AR(p)$ 模型的 p 个特征根都在单位圆内。所以 $AR(p)$ 模型平稳的充要条件是它的 p 个特征根都在单位圆内。

根据特征根和自回归系数多项式的根成倒数的性质，AR 模型平稳的等价判别条件是该 AR 模型的自回归系数多项式的根，即 $\Phi(u)=0$ 的根，都在单位圆外。

2. 平稳域判别

对于一个 $AR(p)$ 模型而言，如果没有平稳性的要求，实际上也就意味着对参数向量 $(\phi_1,\phi_2,\cdots,\phi_p)'$ 没有任何限制，它们可以取遍 p 维欧氏空间的任意一点，但是如果加上了平稳性限制，参数向量 $(\phi_1,\phi_2,\cdots,\phi_p)'$ 就只能取 p 维欧氏空间的一个子集，使得特征根都在单位圆内的系数集合

$$\{\phi_1,\phi_2,\cdots,\phi_p\,|\,\text{单位根都在单位圆内}\}$$

被称为 $AR(p)$ 模型的平稳域。

对于低 AR 阶模型用平稳域的方法判别模型的平稳性通常更为简便。

(1) $AR(1)$ 模型的平稳域。

$AR(1)$ 模型为：$x_t=\phi x_{t-1}+\varepsilon_t$，其特征方程为：$\lambda-\phi=0$，特征根为：$\lambda=\phi$。根据 AR 模型平稳的充要条件，容易推出 $AR(1)$ 模型平稳的充要条件是

$$|\phi|<1 \tag{3.10}$$

所以，$AR(1)$ 模型的平稳域就是 $\{\phi \mid -1<\phi<1\}$。

(2) $AR(2)$ 模型的平稳域。

$AR(2)$ 模型为：$x_t=\phi_1 x_{t-1}+\phi_2 x_{t-2}+\varepsilon_t$，其特征方程为：$\lambda^2-\phi_1\lambda-\phi_2=0$，特征根为：$\lambda_1=\dfrac{\phi_1+\sqrt{\phi_1^2+4\phi_2}}{2}$，$\lambda_2=\dfrac{\phi_1-\sqrt{\phi_1^2+4\phi_2}}{2}$。根据 AR 模型平稳的充要条件，$AR(2)$ 模型平稳的充要条件是 $|\lambda_1|<1$ 且 $|\lambda_2|<1$。

根据一元二次方程的性质和 $AR(2)$ 模型的平稳条件，有

$$\begin{cases}\lambda_1+\lambda_2=\phi_1\\ \lambda_1\lambda_2=-\phi_2\end{cases},\text{且}\ |\lambda_1|<1,|\lambda_2|<1 \tag{3.11}$$

可以导出：

(1) $|\phi_2|=|\lambda_1\lambda_2|<1$

(2) $\phi_2+\phi_1=-\lambda_1\lambda_2+\lambda_1+\lambda_2=1-(1-\lambda_1)(1-\lambda_2)<1$

(3) $\phi_2-\phi_1=-\lambda_1\lambda_2-\lambda_1-\lambda_2=1-(1+\lambda_1)(1+\lambda_2)<1$

这三个限制条件意味着 $AR(2)$ 模型的平稳域是一个三角性区域，见图 3—2。

$\{\phi_1,\phi_2 \mid |\phi_2|<1,\text{且}\ \phi_2\pm\phi_1<\}$

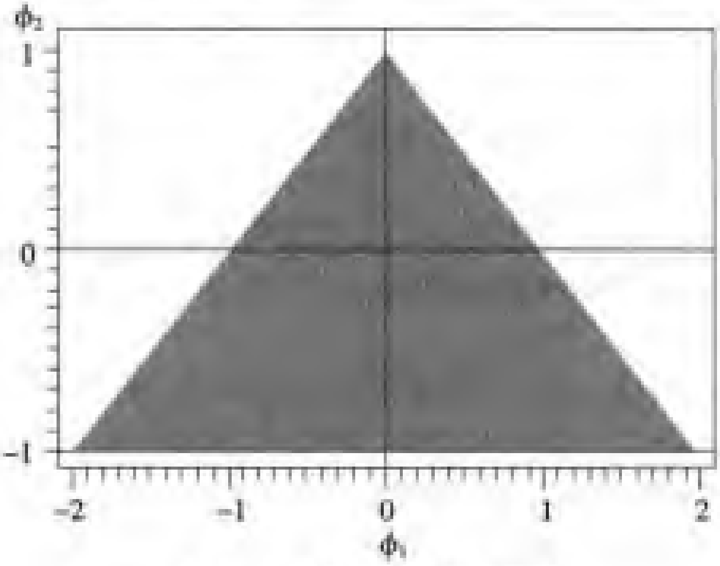


图 3—2 $AR(2)$ 模型的平稳域

【例 3.1 续】 分别用特征根判别法和平稳域判别法检验例 3.1 中四个 AR 模型的平稳性：

(1) $x_t=0.8x_{t-1}+\varepsilon_t$　　(2) $x_t=-1.1x_{t-1}+\varepsilon_t$

(3) $x_t=x_{t-1}-0.5x_{t-2}+\varepsilon_t$　　(4) $x_t=x_{t-1}+0.5x_{t-1}+\varepsilon_t$

其中$\{\varepsilon_t\}$均为服从标准正态分布的白噪声序列。

结论见表 3—1。

表 3—1

模型	特征根判别	平稳域判别	结论
(1)	$\lambda_1=0.8$	$\phi=0.8$	平稳
(2)	$\lambda_2=-1.1$	$\phi=-1.1$	非平稳
(3)	$\lambda_1=\frac{1+i}{2},\lambda_2=\frac{1-i}{2}$	$\|\phi_2\|=0.5,\phi_2+\phi_1=0.5,\phi_2-\phi_1=-1.5$	平稳
(4)	$\lambda_1=\frac{1+\sqrt{3}}{2},\lambda_2=\frac{1-\sqrt{3}}{2}$	$\|\phi_2\|=0.5,\phi_2+\phi_1=1.5,\phi_2-\phi_1=-0.5$	非平稳

显然，理论判别得到的结论支持例 3.1 根据时序图（图 3—1）所得出的直观判断。

三、平稳 AR 模型的统计性质

1. 均值

值如 $AR(p)$模型（3.5）满足了平稳性条件，在等式两边取期望，得

$$Ex_t=E(\phi_0+\phi_1x_{t-1}+\phi_2x_{t-2}+\cdots+\phi_px_{t-p}+\varepsilon_r) \tag{3.12}$$

根据平稳序列均值为常数的性质，有 $Ex_t=\mu$，$\forall_t\in T$，且因为$\{\varepsilon_t\}$为白噪声序列，有 $E\varepsilon_t=0$，所以式（3.12）等价于

$$(1-\phi_1-\cdots-\phi_p)\mu=\phi_0$$

$$\Rightarrow\mu=\frac{\phi_0}{1-\phi_1-\cdots-\phi_p}$$

特别地，对于中心化 $AR(p)$模型，有 $Ex_t=0$。

2. 方差

(1) Green 函数。要得到平稳 $AR(p)$模型的方差，需要借助 Green 函数的帮助，下面给出 Green 函数的定义。

设 λ_1，λ_2，…，λ_p 为平稳 $AR(p)$模型的特征根，则平稳 $AR(p)$模型可以写成如下形式：

$$x_t=\frac{\varepsilon_t}{\Phi(B)}$$

$$= \sum_{i=1}^{p} \frac{k_i}{1-\lambda_i B}\varepsilon_t$$

$$= \sum_{i=1}^{p} \sum_{j=0}^{\infty} k_i (\lambda_i B)^j \varepsilon_t \qquad (3.13)$$

$$= \sum_{j=0}^{\infty} \sum_{i=1}^{p} k_i \lambda_i^j \varepsilon_{t-j}$$

$$\hat{=} \sum_{j=0}^{\infty} G_j \varepsilon_{t-j}$$

式中，$G_j = \sum_{i=1}^{p} k_i \lambda_i^j (j = 1,2,\cdots)$。

式（3.13）称为 AR 模型的传递形式，系数 G_j（$j=1$，2，…）称为 Green 函数。因为 λ_1，λ_2，…，λ_p 都在单位圆内，所以 Green 函数应该呈负指数下降，且 $\lim\limits_{j\to\infty} |G_j| = 0$。通过待定系数法，可以确定 Green 函数的递推公式。

记 $G(B) = \sum_{j=0}^{\infty} G_j B^j$，$AR(p)$ 模型（3.13）可以简记为：

$$x_t = G(B)\varepsilon_t \qquad (3.14)$$

把式（3.14）带入 $AR(p)$ 模型 $\Phi(B)x_t = \varepsilon_t$，得到

$$\Phi(B)G(B)\varepsilon_t = \varepsilon_t$$

展开上式，得

$$(1 - \sum_{k=1}^{p} \phi_k B^k)(1 + \sum_{j=0}^{\infty} G_k B^j)\varepsilon_t = \varepsilon_t$$

整理得

$$\left[1 + \sum_{j=1}^{\infty} (G_j - \sum_{k=1}^{j} \phi'_k G_{j-k}) B^j\right]\varepsilon_t = \varepsilon_t$$

根据待定系数法，有

$$G_j - \sum_{k=1}^{j} \phi'_k G_{j-k} = 0, j = 1,2,\cdots$$

由此可以得到 Green 函数递推公式：

$$\begin{cases} G_0 = 1 \\ G_j = \sum_{k=1}^{j} \phi'_k G_{j-k}, j=1, 2, \cdots \end{cases} \qquad (3.15)$$

式中：

$$\phi'_k = \begin{cases} \phi_k, k \leqslant p \\ 0, k \end{cases}$$

(2) 平稳 AR 模型的方差。对平稳 AR 模型（3.14）等式两边求方差，有

$$\mathrm{Var}(x_t) = \sum_{j=0}^{\infty} G_j^2 \mathrm{Var}(\varepsilon_t)$$

式中，$\{\varepsilon_t\}$为白噪声序列；$\mathrm{Var}(\varepsilon_t)=\sigma_j^2$。所以

$$\mathrm{Var}(x_t) = \sum_{j=0}^{\infty} G_j^2 \sigma_\varepsilon^2 \tag{3.16}$$

因为 G_j 呈负指数下降，所以 $\sum_{j=0}^{\infty} G_j^2 < \infty$，这说明平稳序列 $\{x_t\}$ 方差有界，等于常数 $\sum_{j=0}^{\infty} G_j^2 \sigma_\varepsilon^2$。

【例 3.2】 求平稳 $AR(1)$模型的方差。

把平稳 $AR(1)$模型转化为传递形式：

$$(1-\phi_1 B)x_t = \varepsilon_t$$

$$\Rightarrow x_t = \frac{\varepsilon_t}{1-\phi_1 B} = \sum_{i=0}^{\infty} (\phi_1 B)^i \varepsilon_t = \sum_{i=0}^{\infty} \phi_1^i \varepsilon_{t-i}$$

Green 函数为

$$G_j = \phi_1^j, j=0,1,\cdots$$

所以平稳 $AR(1)$模型的方差为：

$$\mathrm{Var}(x_t) = \sum_{j=0}^{\infty} G_j^2 \mathrm{Var}(\varepsilon_t) = \sum_{j=0}^{\infty} \phi_1^{2j} \sigma_\varepsilon^2 = \frac{\sigma_\varepsilon^2}{1-\phi_1^2}$$

3. 协方差函数

在平稳模型 $x_t=\phi_1 x_{t-1}+\phi_2 x_{t-2}+\cdots+\phi_p x_{t-p}+\varepsilon_t$ 等号两边同乘 x_{t-k}，$\forall k \geq 1$，再求期望，得

$$E(x_t x_{t-k}) = \phi_1 E(x_{t-1} x_{t-k}) + \cdots + \phi_p E(x_{t-p} x_{t-k}) + E(\varepsilon_t x_{t-k}), \forall k \geq 1$$

根据式（3.4）$AR(p)$模型的条件三，有

$$E(\varepsilon_t x_{t-k}), \forall k \geq 1$$

于是可以得到如下自协方差函数的递推公式：

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \cdots + \phi_p \gamma_{k-p} \tag{3.17}$$

【例 3.3】 求平稳 $AR(1)$模型的自协方差函数。

平稳 $AR(1)$模型的自协方差函数递推公式为：

$$\begin{aligned} \gamma_k &= \phi_1 \gamma_{k-1} \\ &= \phi_1^k \gamma_0 \end{aligned}$$

根据例 3.2 已知

$$\gamma_0=\frac{\sigma_\varepsilon^2}{1-\phi_1^2}$$

所以平稳 $AR(1)$模型的自协方差函数递推公式如下：

$$\gamma_k=\phi_1^k\frac{\sigma_\varepsilon^2}{1-\phi_1^2},\quad \forall k\geqslant 1$$

【例 3.4】 求平稳 $AR(2)$模型的自协方差函数。

平稳 $AR(2)$模型的递推公式为：

$$\gamma_k=\phi_1\gamma_{k-1}+\phi_2\gamma_{k-2},\forall k\geqslant 1$$

特别地，当 $k=1$ 时，有

$$\gamma_1=\phi_1\gamma_0+\phi_2\gamma_1$$

即

$$\gamma_1=\frac{\phi_1\gamma_0}{1-\phi_2}$$

利用 Green 函数可以推导出 $AR(2)$模型的方差为：

$$\gamma_0=\frac{1-\phi_2}{(1+\phi_2)(1-\phi_1-\phi_1)(1+\phi_1-\phi_2)}\sigma_\varepsilon^2$$

所以平稳 $AR(2)$模型的自协方差函数的递推公式如下：

$$\begin{cases}\gamma_0=\dfrac{1-\phi_2}{(1+\phi_2)\ (1-\phi_1-\phi_2)\ (1+\phi_1-\phi_2)}\sigma_\varepsilon^2\\ \gamma_1=\dfrac{\phi_1\gamma_0}{1-\phi_2}\\ \gamma_k=\phi_1\gamma_{k-1}+\phi_2\gamma_{k-2},\quad k\geqslant 2\end{cases}$$

4. 自相关系数

(1) 平稳 AR 模型自相关系数递推公式。由于 $\rho_k=\frac{\gamma_k}{\gamma_0}$，在自协方差函数的递推公式（3.17）等号两边同除以方差函数 γ_0，就得到自相关系数的递推公式：

$$\rho_k=\phi_1\rho_{k-1}+\phi_1\rho_{k-2}+\cdots+\phi_p\rho_{k-p} \tag{3.18}$$

容易验证平稳 $AR(1)$模型的自相关系数递推公式为：

$$\rho_k=\phi_1^k,k\geqslant 0$$

平稳 $AR(2)$模型的自相关系数递推公式为：

$$\rho_k\begin{cases}1, & k=0\\ \dfrac{\phi_1}{1-\phi_2}, & k=1\\ \phi_1\rho_{k-1}+\phi_2\rho_{k-2}, & k\geqslant 2\end{cases}$$

(2) 自相关系数的性质。平稳 $AR(p)$模型的自相关系数有两个显著的性质：

一是拖尾性；二是呈负指数衰减。这两个性质都可以从自相关系数的通解推出。

根据式（3.18），容易看出 $AR(p)$ 模型的自相关系数的表达式实际上是一个 p 阶齐次差分方程。那么滞后任意 k 阶的自相关系数的通解为：

$$\rho(k)=\sum_{i=1}^{p}c_i\lambda_i^k \tag{3.19}$$

式中，$|\lambda_i|<1$（$i=1$，…，p）为该差分方程的特征根；c_1，…，c_p 为任意常数。如果已知任意 p 个自相关系数的值就可以确定 c_1，…，c_p 的解，显然 c_1，…，c_p 不能全为零。

通过这个通解形式，容易推出 $\rho(k)$ 始终有非零取值，不会在 k 大于某个常数之后就恒等于零，这个性质就是拖尾性。

可以直观地解释 $AR(p)$ 模型自相关系数拖尾的原因。对于一个平稳 $AR(p)$ 模型：

$$x_t=\phi_1x_{t-1}+\phi_2x_{t-2}+\cdots+\phi_px_{t-p}+\varepsilon_t$$

虽然它的表达式显示 x_t 只受当期随机误差 ε_t 和最近 p 期的序列值 x_{t-1}，…，x_{t-p} 的影响，但是由于 x_{t-1} 的值又依赖于 x_{t-1-p}，所以实际上 x_{t-1-p} 对 x_t 也有影响，依此类推，x_t 之前的每一个序列值 x_{t-1}，…，x_i 都会对 x_t 构成影响。自回归模型的这种特性体现在自相关系数上就是自相关系数的拖尾性。

同时随着时间的推移，$\rho(k)$ 会迅速衰减，因为 $|\lambda_i|<1$（$i=1$，…，p），所以 $k\to\infty$ 时，$\lambda_i^k\to0$（$i=1$，…，p），继而导致 $\rho(k)=\sum_{i=1}^{p}c_i\lambda_i^k\to0$，而且这种影响是以负指数 λ^k 的速度在减小。

这种自相关系数以负指数衰减的性质就是在第 2 章利用自相关图判断平稳序列时所说的“短期相关”性。它是平稳序列的一个重要特征。这个特征表明对平稳序列而言通常只有近期的序列值对现时值的影响比较明显，间隔越远的过去值对现时值的影响越小。

【例 3.5】 考察如下四个平稳 AR 模型的自相关图。

（1）$x_t=0.8x_{t-1}+\varepsilon_t$　　（2）$x_t=-0.8x_{t-1}+\varepsilon_t$

（3）$x_t=x_{t-1}-0.5x_{t-2}+\varepsilon_t$　　（4）$x_t=-x_{t-1}-0.5x_{t-2}+\varepsilon_t$

假定 $\{\varepsilon_t\}$ 为标准正态白噪声序列，拟合这四个 AR 模型，得到样本自相关图如图 3—3 所示（纵轴为自相关系数，横轴为延迟阶数）。

从图 3—3 中我们看到，这四个平稳 AR 模型，不论它们是 $AR(1)$ 模型还是 $AR(2)$ 模型，不论它们的特征根是实根还是复根，是正根还是负根，它们的自相关系数都呈现出拖尾性和呈负指数衰减到零值附近的性质。

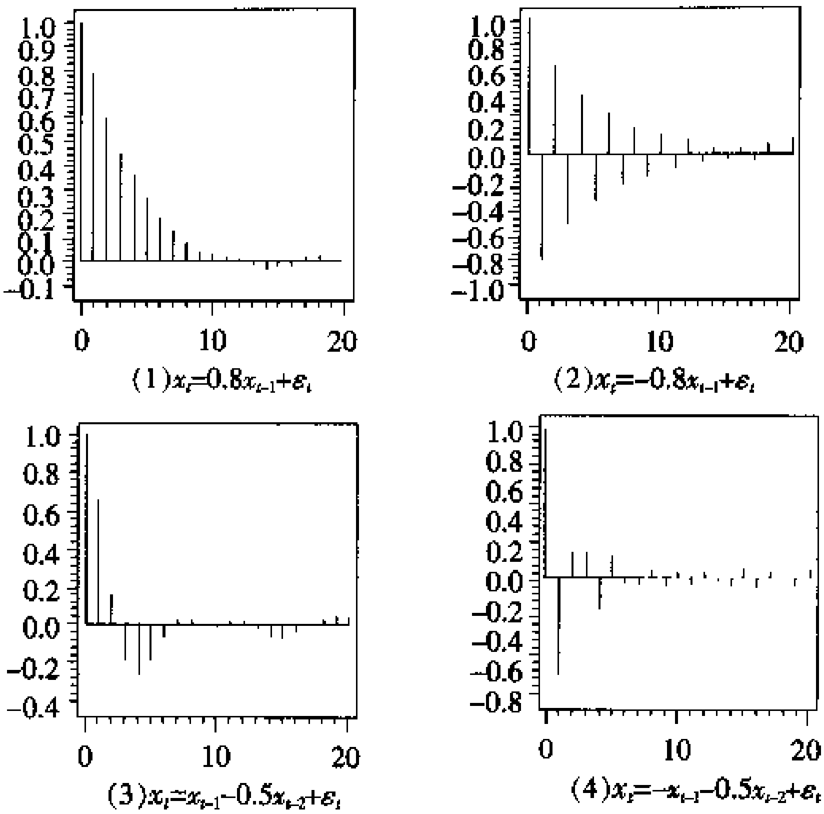


图 3—3 *AR* 模型样本自相关图

但由于特征根的不同，它们的自相关系数衰减的方式也不一样。有的自相关系数是按负指数单调收敛到零（如模型 1）；有的是呈现正负相间地衰减（如模型 2）；还有些自回归系数会呈现出类似于周期性的余弦衰减，即具有“伪周期”特征（如模型 3），这些都是平稳模型自相关系数常见的特征。

5. 偏自相关系数

（1）偏自相关系数的定义。对于一个平稳 $AR(p)$模型，求出滞后 k 自相关系数$\rho(k)$时，实际上得到的并不是 x_t 与 x_{t-k}之间单纯的相关关系。因为 x_t 同时还会受到中间 $k-1$ 个随机变量 $x_{t-1},x_{t-1},\cdots,x_{t-k+1}$的影响，而这 $k-1$ 个随机变量又都和 x_{t-k}具有相关关系，所以自相关系数 $\rho(k)$里实际上掺杂了其他变量对 x_t 与 x_{t-k}的相关影响。为了能单纯测度 x_{t-k}对 x_t 的影响，引进偏自相关系数的概念。

定义 3.3 对于平稳序列$\{x_t\}$，所谓滞后 k 偏自相关系数就是指在给定中间 $k-1$ 个随机变量 $x_{t-1},x_{t-2},\cdots,x_{t-k+1}$的条件下，或者说，在剔除了中间 $k-1$ 个随机变量 $x_{t-1},x_{t-2},\cdots,x_{t-k+1}$的干扰之后，$x_{t-k}$对 x_t 影响的相关度量。用数学语言描述就是：

$$\rho_{x_t,x_{t-k}\mid x_{t-1},\cdots,x_{t-k+1}}=\frac{E[(x_t-\hat{E}x_t)(x_{t-k}-\hat{E}x_{t-k})]}{E[(x_{t-k}-\hat{E}x_{t-k})^2]} \tag{3.20}$$

这就是滞后 k 偏自相关系数的定义。

(2) 偏自相关系数的计算。偏自相关系数的定义和回归分析中偏相关系数的定义非常相似。这可以启发我们从线性回归的角度，得到偏自相关系数的另一层含义。

假定 $\{x_t\}$ 为中心化平稳序列，用过去的 k 期序列值 x_{t-1}，x_{t-2}，…，x_{t-k} 对 x_t 作 k 阶自回归拟合，即

$$x_t=\phi_{k1}x_{t-1}+\phi_{k2}x_{t-2}+\cdots+\phi_{kk}x_{t-k}+\varepsilon_t \tag{3.21}$$

式中，$E(\varepsilon_t)=0$，$E\varepsilon_t x_s=0$，$\forall s<t$。对 x_{t-1}，x_{t-2}，…，x_{t-k+1} 取条件，记

$$\hat{E}x_t=E[x_t|x_{t-1},\cdots,x_{t-k+1}],\hat{E}x_{t-k}=E[x_{t-k}|x_{t-1},\cdots,x_{t-k+1}]$$

则

$$\begin{aligned}\hat{E}x_t=&\phi_{k1}x_{t-1}+\phi_{k2}x_{t-2}+\cdots+\phi_{kk-1}x_{t-k+1}+\phi_{kk}\hat{E}(x_{t-k})\\&+E(\varepsilon_t|x_{t-1},\cdots,x_{t-k+1})\end{aligned} \tag{3.22}$$

已知 $E(\varepsilon_t)=0$，$E\varepsilon_t x_s=0$，$\forall s<t$，所以

$$E(\varepsilon_t|x_{t-1},\cdots,x_{t-k+1})=E(\varepsilon_t)=0$$

式 (3.22) 等价于

$$\hat{E}x_t=\phi_{k1}x_{t-1}+\phi_{k2}x_{t-2}+\cdots+\phi_{kk-1}x_{t-k+1}+\phi_{kk}\hat{E}(x_{t-k})$$

则

$$x_t-\hat{E}x_t=\phi_{kk}(x_{t-k}-\hat{E}x_{t-k})+\varepsilon_t \tag{3.23}$$

在式 (3.23) 等号两边同时乘以 $x_{t-k}-\hat{E}x_{t-k}$，并求期望：

$$\begin{aligned}E[(x_t-\hat{E}x_t)(x_{t-k}-\hat{E}x_{t-k})]=&\phi_{kk}E[(x_{t-k}-\hat{E}x_{t-k})^2]\\&+E[\varepsilon_t(x_{t-k}-\hat{E}x_{t-k})]\end{aligned} \tag{3.24}$$

因为 $E\varepsilon_t x_s=0$，$\forall s<t$，所以

$$E[\varepsilon_t(x_{t-k}-\hat{E}x_{t-k})]=0$$

式 (3.24) 等价于

$$E[(x_t-\hat{E}x_t)(x_{t-k}-\hat{E}x_{t-k})]=\phi_{kk}E[(x_{t-k}-\hat{E}x_{t-k})^2]$$

由此得出

$$\phi_{kk}=\frac{E[(x_t-\hat{E}x_t)(x_{t-k}-\hat{E}x_{kt})]}{E[(x_{t-k}-\hat{E}x_{t-k})^2]} \tag{3.25}$$

式 (3.25) 等号右边的结果正好等于式 (3.20) 所定义的滞后 k 偏自相关系数。

这说明滞后 k 偏自相关系数实际上就等于 k 阶自回归模型第 k 个回归系数 ϕ_{kk} 的值。根据这个性质容易计算偏自相关系数的值。

在式（3.21）等号两边同乘 x_{t-k}，并求期望，得

$$\rho_l=\phi_{k1}\rho_{l-1}+\phi_{k2}\rho_{l-2}+\cdots+\phi_{kk}\rho_{l-k},\ \forall\, l\geqslant 1$$

取前 k 个方程构成的方程组：

$$\begin{cases}\rho_1=\phi_{k1}\rho_0+\phi_{k2}\rho_1+\cdots+\phi_{kk}\rho_{k-1}\\ \rho_2=\phi_{k1}\rho_1+\phi_{k2}\rho_0+\cdots+\phi_{kk}\rho_{k-2}\\ \quad\cdots\cdots\cdots\cdots\\ \rho_k=\phi_{k1}\rho_{k-1}+\phi_{k2}\rho_{k-2}+\cdots+\phi_{kk}\rho_0\end{cases}\tag{3.26}$$

该方程组称为 Yule-Walker 方程。通过解该方程组，可以得到参数 $(\phi_{k1},\phi_{k2},\cdots,\phi_{kk})'$ 的解，参数向量中最后一个参数的解，即为滞后 k 偏自相关系数 ϕ_{kk} 的值。

用矩阵形式表示为：

$$\begin{pmatrix}1 & \rho_1 & \cdots & \rho_{k-1}\\ \rho_1 & 1 & \cdots & \rho_{k-2}\\ \vdots & \vdots & & \vdots\\ \rho_{k-1} & \rho_{k-2} & \cdots & 1\end{pmatrix}\begin{pmatrix}\phi_{k1}\\ \phi_{k2}\\ \vdots\\ \phi_{kk}\end{pmatrix}=\begin{pmatrix}\rho_1\\ \rho_2\\ \vdots\\ \rho_k\end{pmatrix}\tag{3.27}$$

根据线性方程组求解的 Cramer 法则，有

$$\phi_{kk}=\frac{D_k}{D}\tag{3.28}$$

式中：

$$D=\begin{vmatrix}1 & \rho_1 & \cdots & \rho_{k-1}\\ \rho_1 & 1 & \cdots & \rho_{k-2}\\ \vdots & \vdots & & \vdots\\ \rho_{k-1} & \rho_{k-2} & \cdots & 1\end{vmatrix},D_k=\begin{vmatrix}1 & \rho_1 & \cdots & \rho_1\\ \rho_1 & 1 & \cdots & \rho_2\\ \vdots & \vdots & & \vdots\\ \rho_{k-1} & \rho_{k-2} & \cdots & \rho_k\end{vmatrix}$$

D 为式（3.27）中系数矩阵的行列式，D_k 是把 D 中的第 k 个列向量换成式（3.27）等号右边的自相关系数向量后构成的行列式。

（3）偏自相关系数的截尾性。可以证明：平稳 $AR(p)$ 模型的偏自相关系数具有 p 步截尾性。所谓 p 步截尾是指，$\phi_{kk}=0$，$\forall k>p$。要证明这一点实际上只要能证明当 $k>p$ 时，$D_k=0$ 就行。

证明：对任一 $AR(p)$ 模型：

$$x_t=\phi_1x_{t-1}+\phi_2x_{t-2}+\cdots+\phi_px_{t-p}+\varepsilon_t,\quad \forall k>p$$

有如下 Yule-Walker 方程成立：

$$\begin{pmatrix}1 & \rho_1 & \cdots & \rho_{p-1}\\ \rho_1 & 1 & \cdots & \rho_{p-2}\\ \vdots & \vdots & & \vdots\\ \rho_{k-1} & \rho_{k-2} & \cdots & \rho_{k-p}\end{pmatrix}\begin{pmatrix}\phi_1\\ \phi_2\\ \vdots\\ \phi_k\end{pmatrix}=\begin{pmatrix}\rho_1\\ \rho_2\\ \vdots\\ \rho_k\end{pmatrix}\tag{3.29}$$

记 ξ_i（$i=1, 2, \cdots, p$）为式（3.29）系数矩阵中 p 个列向量，η 为式（3.29）中等号右边的自相关系数向量，即

$$\xi_i=\begin{pmatrix}\rho_1\\ \rho_{i-1}\\ \vdots\\ \rho_{i-k}\end{pmatrix},i=1,2,\cdots,p;\eta=\begin{pmatrix}\rho_1\\ \rho_2\\ \vdots\\ \rho_k\end{pmatrix}$$

则有

$$\eta=\phi_1\xi_1+\phi_2\xi_2+\cdots+\phi_p\xi_p \tag{3.30}$$

因为 $AR(p)$模型限制条件之一是 $\phi_p\neq0$，所以向量 η 一定可以表示成向量 ξ_i（$i=1, 2, \cdots, p$)的非零线性组合。

当 $k>p$ 时

$$D_k=\begin{vmatrix}1 & \rho_1 & \cdots & \rho_{p-1} & \cdots & \rho_1\\ \rho_1 & 1 & \cdots & \rho_{p-2} & \cdots & \rho_2\\ \vdots & \vdots & & \vdots & & \vdots\\ \rho_{k-1} & \rho_{k-2} & \cdots & \rho_{k-p} & \cdots & \rho_k\end{vmatrix} \tag{3.31}$$

显然 D_k 的前 p 个列向量正好就是 ξ_i（$i=1, 2, \cdots, p$），而最后一个列向量正好就是向量 η。根据式（3.30），说明在行列式（3.31）中最后一个列向量可以用另外 p 个列向量的线性组合表示。根据行列式的性质，具有这种线性相关关系的行列式的值一定为零，即 $D_k=0$。由 $D_k=0$，等价推出 $\phi_{kk}=0$。

证毕。

由此证明了 $AR(p)$模型偏自相关系数的 p 步截尾性。这个性质连同前面的自相关系数拖尾性是 $AR(p)$模型重要的识别依据。

【例 3.5 续】 考察例 3.5 中四个平稳 AR 模型的偏自相关系数截尾性。

（1）$x_t=0.8x_{t-1}+\varepsilon_t$　　（2）$x_t=-0.8x_{t-1}+\varepsilon_t$

（3）$x_t=x_{t-1}-0.5x_{t-2}+\varepsilon_t$　　（4）$x_t=-x_{t-1}-0.5x_{t-2}+\varepsilon_t$

根据公式（3.28）容易算出，$AR(1)$模型的偏自相关系数为：

$$\phi_{kk}=\begin{cases}\phi_1, & k=1\\ 0, & k\geq2\end{cases}$$

$AR(2)$模型的偏自相关系数为：

$$\phi_{kk}=\begin{cases}\dfrac{\phi_1}{1-\phi_2}, & k=1\\ \phi_2, & k=2\\ 0, & k\geq3\end{cases}$$

所以，这四个 AR 模型的理论偏自相关系数如表 3—2 所示。

表 3—2

(1) $x_t=0.8x_{t-1}+\varepsilon_t$	$\phi_{kk}=\begin{cases}0.8, & k=1\\0, & k\geq2\end{cases}$
(2) $x_t=-0.8x_{t-1}+\varepsilon_t$	$\phi_{kk}=\begin{cases}-0.8, & k=1\\0, & k\geq2\end{cases}$
(3) $x_t=x_{t-1}-0.5x_{t-2}+\varepsilon_t$	$\phi_{kk}=\begin{cases}\frac{2}{3}, & k=1\\-0.5, & k=2\\0, & k\geq3\end{cases}$
(4) $x_t=-x_{t-1}-0.5x_{t-2}+\varepsilon_t$	$\phi_{kk}=\begin{cases}-\frac{2}{3}, & k=1\\-0.5, & k=2\\0, & k\geq3\end{cases}$

假定$\{\varepsilon_t\}$为标准正态白噪声序列，拟合这四个 AR 模型，得到样本偏自相关图如图 3—4 所示（纵轴为偏自相关系数，横轴为延迟阶数）。

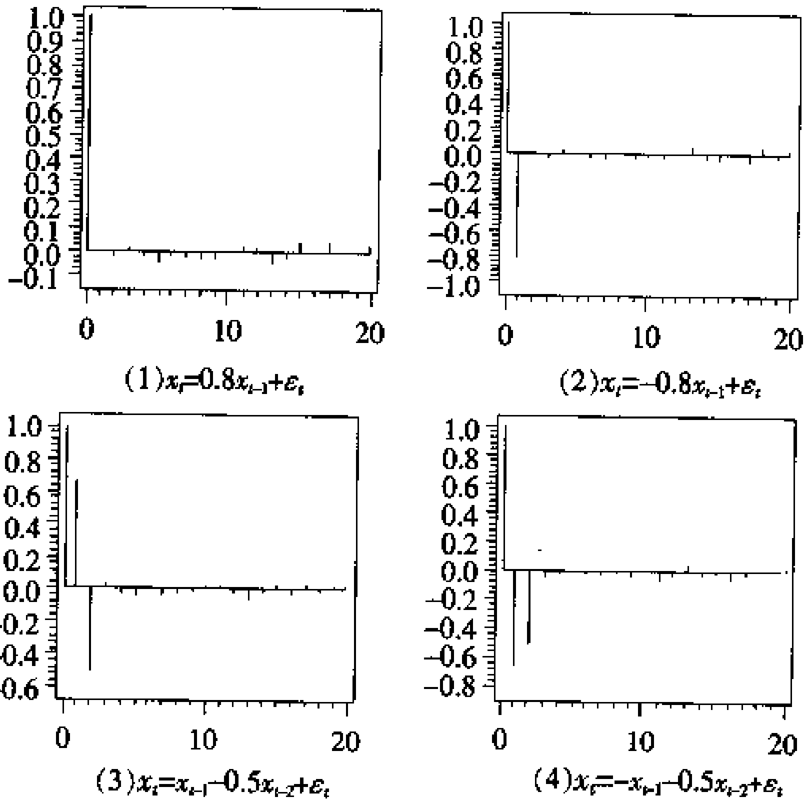


图 3—4 AR 模型样本偏自相关图

由于样本的随机性，样本偏自相关系数不会和理论偏相关系数一样严格截尾，但可以看出两个 $AR(1)$模型的样本偏相关系数 1 阶显著不为零，1 阶之后都近似为零；而两个 $AR(2)$模型的样本偏相关系数 2 阶显著不为零，2 阶之后

都近似为零。样本偏自相关图可以直观地验证 AR 模型偏自相关系数截尾性。

3.2.2 *MA* 模型

一、定义

定义 3.4 具有如下结构的模型称为 q 阶移动平均（moving average）模型，简记为 $MA(q)$：

$$\begin{cases} x_t=\mu+\varepsilon_t-\theta_1\varepsilon_{t-1}-\theta_2\varepsilon_{t-2}-\cdots-\theta_q\varepsilon_{t-q} \\ \theta_q\neq 0 \\ E(\varepsilon_t)=0,\mathrm{Var}(\varepsilon_t)=\sigma_\varepsilon^2,E(\varepsilon_t\varepsilon_s)=0,s\neq t \end{cases} \tag{3.32}$$

使用 $MA(q)$ 模型需要满足两个限制条件：

条件一：$\theta_q\neq 0$，这个限制条件保证了模型的最高阶数为 q。

条件二：$E(\varepsilon_t)=0$，$\mathrm{Var}(\varepsilon_t)=\sigma_\varepsilon^2$，$E(\varepsilon_t\varepsilon_s)=0$，$s\neq t$。即随机干扰序列$\{\varepsilon_t\}$为零均值白噪声序列

$$\varepsilon_t\sim WN(0,\sigma_\varepsilon^2)$$

通常缺省默认式（3.32）的限制条件，把模型简记为：

$$x_t=\mu+\varepsilon_t-\theta_1\varepsilon_{t-1}-\theta_2\varepsilon_{t-2}-\cdots-\theta_q\varepsilon_{t-q} \tag{3.33}$$

当时 $\mu=0$，模型（3.32）移为中心化 $MA(q)$ 模型。对非中心化 $MA(q)$ 模型只要做一个简单的位移 $y_t=x_t-\mu$，就可以转化为中心化 $MA(q)$ 模型。这种中心化运算不会影响序列值之间的相关关系，所以今后在分析 MA 模型的相关关系时，常常简化为对它的中心化模型进行分析。

使用延迟算子，中心化 $MA(q)$ 模型又可以简记为：

$$x_t=\Theta(B)\varepsilon_t$$

式中，$\Theta(B)=1-\theta_1B-\theta_2B^2-\cdots-\theta_qB^q$，称为 q 阶移动平均系数多项式。

二、*MA* 模型的统计性质

1. 常数均值

当 $q<\infty$ 时，$MA(q)$ 模型具有常数均值

$$Ex_t=E(\mu+\varepsilon_t-\theta_1\varepsilon_{t-1}-\theta_2\varepsilon_{t-2}-\cdots-\theta_q\varepsilon_{t-q})=\mu$$

特别地，如果该模型为中心化 $MA(q)$ 模型，该模型均值为零。

2. 常数方差

$$\mathrm{Var}(x_t)=\mathrm{Var}(\mu+\varepsilon_t-\theta_1\varepsilon_{t-1}-\theta_2\varepsilon_{t-2}-\cdots-\theta_q\varepsilon_{t-q})=(1+\theta_1^2+\cdots+\theta_q^2)\sigma_\varepsilon^2$$

3. 自协方差函数只与滞后阶数相关，且 q 阶截尾

$$\begin{aligned}\gamma_k &= E(x_tx_{t-k}) \\ &= E[(\mu+\varepsilon_t-\theta_1\varepsilon_{t-1}-\cdots-\theta_q\varepsilon_{t-q})(\mu+\varepsilon_{t-k}-\theta_1\varepsilon_{t-k-1}-\cdots-\theta_q\varepsilon_{t-k-q})]\end{aligned}$$

$$
=\begin{cases}(1+\theta_1^2+\cdots+\theta_q^2)\sigma_\varepsilon^2, & k=0\\ (-\theta_k+\sum_{i=1}^{q-k}\theta_i\theta_{k+i})\sigma_\varepsilon^2, & 1\leqslant k\leqslant q\\ 0, & k>q\end{cases}
$$

4. 自相关系数 q 阶截尾

$$
\rho_k=\frac{\gamma_k}{\gamma_0}=\begin{cases}1, & k=0\\ \dfrac{-\theta_k+\sum_{i=1}^{q-k}\theta_i\theta_{k+i}}{1+\theta_1^2+\cdots+\theta_q^2}, & 1\leqslant k\leqslant q\\ 0, & k>q\end{cases}
$$

容易验证，$MA(1)$模型的自相关系数为：

$$
\rho_k=\begin{cases}1, & k=0\\ \dfrac{-\theta_1}{1+\theta_1^2}, & k=1\\ 0, & k\geqslant 2\end{cases}
$$

$MA(2)$模型的自相关系数为：

$$
\rho_k=\begin{cases}1, & k=0\\ \dfrac{-\theta_1+\theta_1\theta_2}{1+\theta_1^2+\theta_2^2}, & k=1\\ \dfrac{-\theta_2}{1+\theta_1^2+\theta_2^2}, & k=2\\ 0, & k\geqslant 3\end{cases}
$$

5. 偏自相关系数拖尾

仍然借用 AR 模型中滞后 k 偏自相关系数的记号 ϕ_{kk} 来表示 $MA(q)$模型的滞后 k 偏自相关系数。容易证明它拖尾。

证明：在中心化 $MA(q)$模型场合，有

$$
\phi_{kk}=\frac{E(x_t x_{t-k}\mid x_{t-1},\cdots,x_{t-k-1})}{\mathrm{Var}(x_{t-k}\mid x_{t-1},\cdots,x_{t-k-1})}
$$

$$
=\frac{E[(\varepsilon_t-\theta_1\varepsilon_{t-1}-\cdots-\theta_q\varepsilon_{t-q})(\varepsilon_{t-k}-\theta_1\varepsilon_{t-k-1}-\cdots-\theta_q\varepsilon_{t-k-q})\mid\varepsilon_{t-1},\cdots,\varepsilon_{t-k-q+1}]}{\mathrm{Var}(\varepsilon_{t-k}-\theta_1\varepsilon_{t-k-1}-\cdots-\theta_q\varepsilon_{t-k-1}\mid\varepsilon_{t-1},\cdots,\varepsilon_{t-k-q+1})}
$$

$$
=(-\theta_1\varepsilon_{t-1}-\cdots-\theta_q\varepsilon_{t-q})(-\theta_1\varepsilon_{t-k-1}-\cdots-\theta_q\varepsilon_{t-k-q+1})
$$

在 $MA(q)$模型场合，x_{t-1}，…，x_{t-k-1}给定，就等价于 ε_{t-1}，…，$\varepsilon_{t-k-q+1}$给定，该模型滞后 k 偏自相关系数 ϕ_{kk} 就是 ε_{t-1}，…，$\varepsilon_{t-k-q+1}$的函数，ε_{t-1}，…，$\varepsilon_{t-k-q+1}$不会恒等于零，θ_1，…，θ_q 又是任意给定的，所以 ϕ_{kk} 不可能在有限阶之后恒为零。所以 $MA(q)$模型的偏自相关系数拖尾。

证毕。

从上面的分析中，可以得到如下两个重要结论：

(1) 当 $q<\infty$时，$MA(q)$模型一定为平稳模型。

(2) $MA(q)$模型的偏自相关系数拖尾，自相关系数 q 阶截尾性。

$MA(q)$模型的这个性质和 $AR(p)$模型的自相关系数拖尾，偏自相关系数 p 阶截尾的性质正好呈对偶关系。这两个性质非常重要，它们都将在模型拟合时起关键作用。

【例 3.6】 绘制下列 MA 模型的自相关系数图和偏自相关系数图，直观考察 MA 模型自相关系数截尾和偏自相关系数拖尾性。

(1) $x_t=\varepsilon_t-2\varepsilon_{t-1}$ (2) $x_t=\varepsilon_t-0.5\varepsilon_{t-1}$

(3) $x_t=\varepsilon_t-\frac{4}{5}\varepsilon_{t-1}+\frac{16}{25}\varepsilon_{t-2}$ (4) $x_t=\varepsilon_t-\frac{5}{4}\varepsilon_{t-1}+\frac{25}{16}\varepsilon_{t-2}$

假定$\{\varepsilon_t\}$为标准正态白噪声序列。

首先考察自相关系数的特征，见图 3—5（纵轴为样本自相关系数，横轴为延迟阶数）。

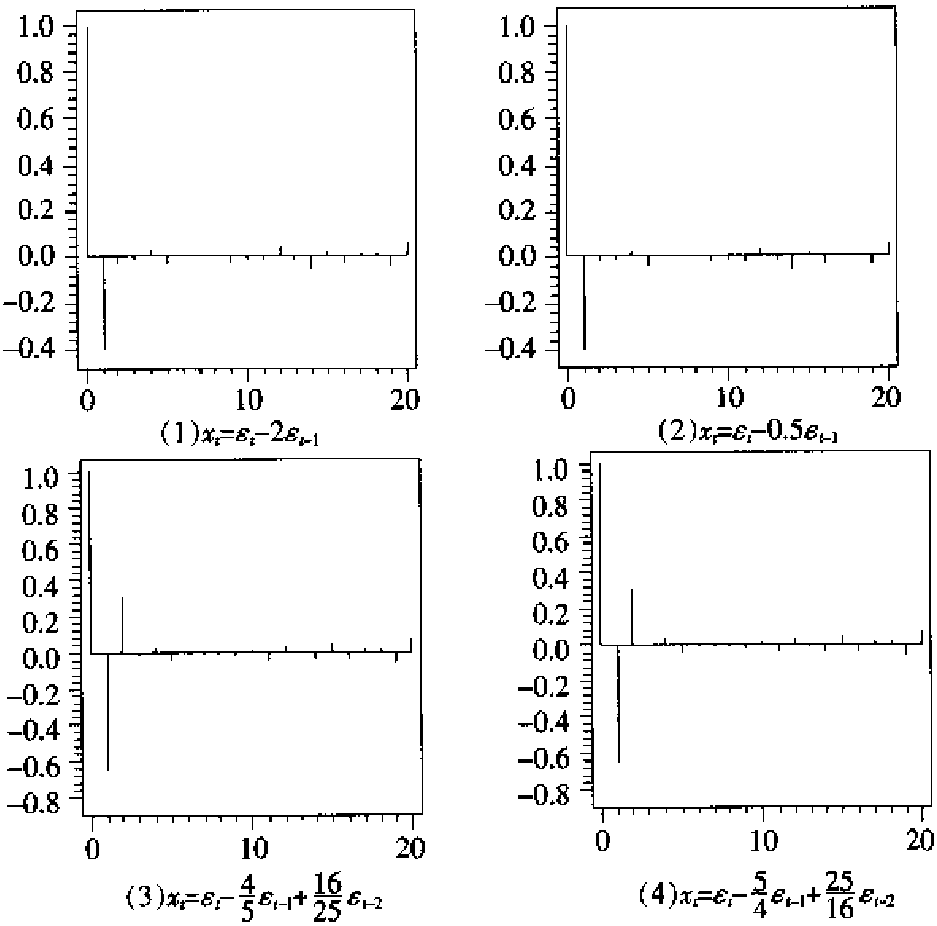
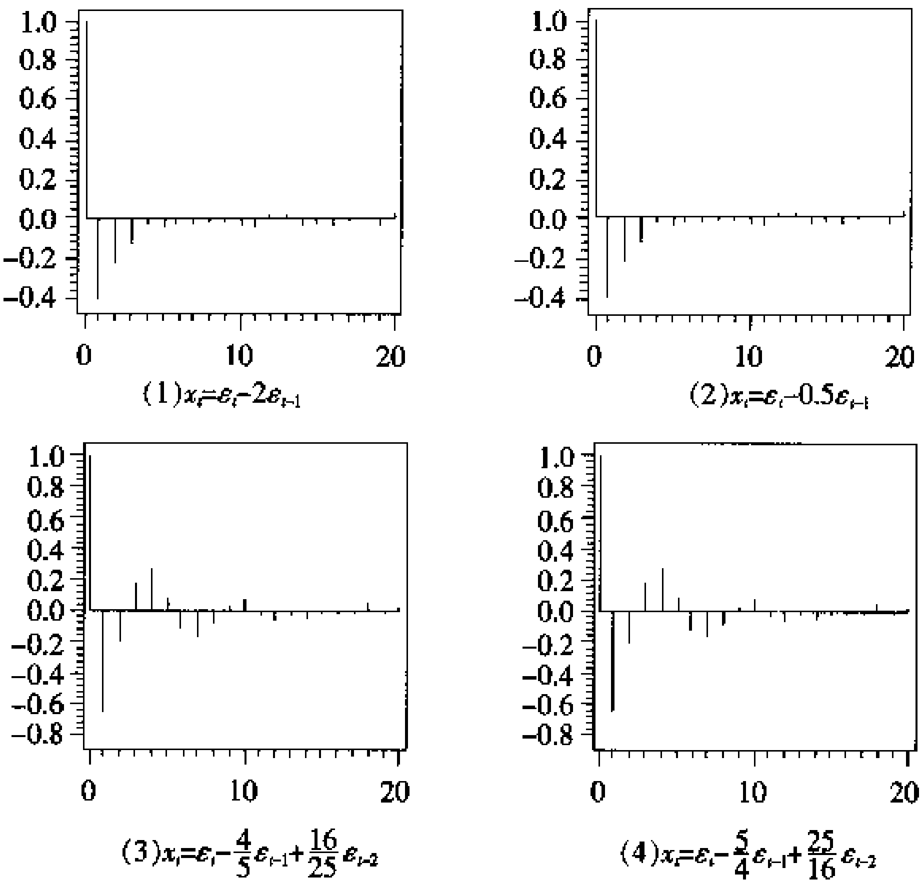


图 3—5 MA 模型的样本自相关图

排除样本随机性的影响，样本自相关图清晰显示出 $MA(1)$ 模型自相关系数 1 阶截尾，$MA(2)$ 模型自相关系数 2 阶截尾的特性。

再考察偏自相关系数的拖尾性质，见图 3—6（纵轴为样本偏自相关系数，横轴为延迟阶数）。



(1) $x_t=\varepsilon_t-2\varepsilon_{t-1}$

(2) $x_t=\varepsilon_t-0.5\varepsilon_{t-1}$

(3) $x_t=\varepsilon_t-\frac{4}{5}\varepsilon_{t-1}+\frac{16}{25}\varepsilon_{t-2}$

(4) $x_t=\varepsilon_t-\frac{5}{4}\varepsilon_{t-1}+\frac{25}{16}\varepsilon_{t-2}$

图 3—6 MA 模型的样本偏自相关图

三、MA 模型的可逆性

再次观察例 3.6 中四个 MA 模型的自相关图（图 3—5），可以发现两个不同的 $MA(1)$ 模型：

(1) $x_t=\varepsilon_t-2\varepsilon_{t-1}$

(2) $x_t=\varepsilon_t-0.5\varepsilon_{t-1}$

具有一模一样的样本自相关图。容易验证它们的理论自相关系数也正好相等：

$$\rho_k=\begin{cases}-0.4, & k=1\\ 0, & k\geq 2\end{cases}$$

而另两个 $MA(2)$ 模型：

(3) $x_t=\varepsilon_t-\frac{4}{5}\varepsilon_{t-1}+\frac{16}{25}\varepsilon_{t-2}$

(4)$x_t=\varepsilon_t-\frac{5}{4}\varepsilon_{t-1}+\frac{25}{16}\varepsilon_{t-2}$

也出现了同样的问题，不同的模型却拥有完全相同的自相关系数：

$$\rho_k=\begin{cases}-0.640\ 12, & k=1\\ 0.312\ 256, & k=2\\ 0, & k\geqslant 3\end{cases}$$

产生这种现象的原因就是我们在第 2 章中提到的：自相关系数有可能不惟一。

这种自相关系数的不惟一性会给我们将来的工作增加麻烦。因为，将来我们都是通过样本自相关系数显示出来的特征选择合适的模型拟合序列的发展，如果自相关系数和模型之间不是一一对应关系，就将导致拟合模型和随机序列之间不会是一一对应关系。

为了保证一个给定的自相关函数能够对应惟一的 *MA* 模型，我们就要给模型增加约束条件。这个约束条件称为 *MA* 模型的可逆性条件。

1. 可逆的定义

容易验证当两个 *MA*(1)模型具有如下结构时，它们的自相关系数正好相等：

模型 1：$x_t=\varepsilon_t-\theta\varepsilon_{t-1}$　　　模型 2：$x_t=\varepsilon_t-\frac{1}{\theta}\varepsilon_{t-1}$

把这两个 *MA*(1)模型表示成两个自相关模型形式：

模型 1：$\frac{x_t}{1-\theta B}=\varepsilon_t$　　　模型 2：$\frac{x_t}{1-\frac{1}{\theta}B}=\varepsilon_t$

显然，$|\theta|<1$ 时，模型 1 收敛，而模型 2 不收敛；$|\theta|>1$ 时，则模型 2 收敛，而模型 1 不收敛。若一个 *MA* 模型能够表示成收敛的 *AR* 模型形式，那么该 *MA* 模型称为可逆模型。一个自相关系数惟一对应一个可逆 *MA* 模型。

2. *MA*(*q*)模型的可逆性条件

与分析 *AR*(*p*)模型的平稳性条件完全类似，*MA*(*q*)模型可以表示为：

$$\varepsilon_t=\frac{x_t}{\Theta(B)} \tag{3.34}$$

式中，$\Theta(B)=1-\theta_1B-\cdots-\theta_qB^q$ 为移动平均系数多项式。假定$\frac{1}{\lambda_1}$，…，$\frac{1}{\lambda_q}$是该系数多项式的 q 个根，则 $\Theta(B)$可以分解成：

$$\Theta(B)=\prod_{k=1}^{q}(1-\lambda_kB) \tag{3.35}$$

把式（3.35）带入式（3.34），得

$$\varepsilon_t=\frac{x_t}{(1-\lambda_1 B)\cdots(1-\lambda_q B)} \tag{3.36}$$

式（3.36）收敛的充要条件是：$|\lambda_i|<1$，等价于 $MA(q)$ 模型的系数多项式的根都在单位圆外 $\left|\frac{1}{\lambda_i}\right|>1$。这个条件也称为 $MA(q)$ 模型的可逆性条件。

显然，$MA(q)$ 模型的可逆概念和 $AR(p)$ 模型的平稳概念是完全对偶的概念。

3. 逆函数的递推公式

如果一个 $MA(q)$ 模型满足可逆性条件，它就可以写成如下两种等价形式：

$$\begin{cases}\Theta(B)\varepsilon_t=x_t & \text{(a)}\\ \varepsilon_t=I(B)x_t & \text{(b)}\end{cases}$$

把式（b）带入式（a），得

$$\Theta(B)I(B)x_t=x_t$$

和 Green 函数的递推公式完全雷同，由待定系数法容易得到逆函数的递推公式为：

$$\begin{cases}I_0=1\\ I_l=\sum_{j=1}^{l}\theta'_j I_{l-j},\ l\geqslant 1\end{cases} \tag{3.37}$$

式中：

$$\theta'_i=\begin{cases}\theta_i,i\leqslant q\\ 0,i>q\end{cases}$$

【例 3.6 续】 考察例 3.6 中四个 MA 模型的可逆性，并写出可逆 MA 模型的逆转形式。

（1）$x_t=\varepsilon_t-2\varepsilon_{t-1}$

（2）$x_t=\varepsilon_t-0.5\varepsilon_{t-1}$

（3）$x_t=\varepsilon_t-\frac{4}{5}\varepsilon_{t-1}+\frac{16}{25}\varepsilon_{t-2}$

（4）$x_t=\varepsilon_t-\frac{5}{4}\varepsilon_{t-1}+\frac{25}{16}\varepsilon_{t-2}$

见表 3—3。

表 3—3

模型	条件	结论
（1）$x_t=\varepsilon_t-2\varepsilon_{t-1}$	$\lvert\theta_1\rvert=2>1$	不可逆
（2）$x_t=\varepsilon_t-0.5\varepsilon_{t-1}$	$\lvert\theta_1\rvert=0.5<1$	可逆
（3）$x_t=\varepsilon_t-\frac{4}{5}\varepsilon_{t-1}+\frac{16}{25}\varepsilon_{t-2}$	$\theta_2=\frac{16}{25}<1$ $\theta_2+\theta_1=-\frac{16}{25}+\frac{4}{5}=\frac{4}{25}<1$ $\theta_2-\theta_1=-\frac{16}{25}-\frac{4}{5}=-\frac{36}{25}<1$	可逆
（4）$x_t=\varepsilon_t-\frac{5}{4}\varepsilon_{t-1}+\frac{25}{16}\varepsilon_{t-2}$	$\theta_2=\frac{25}{16}>1$	不可逆

模型 2：$x_t=\varepsilon_t-0.5\varepsilon_{t-1}$的逆函数为：

$$\begin{cases} I_0=1 \\ I_l=0.5^l, l\geq 1 \end{cases}$$

该模型的逆转形式为：

$$\varepsilon_t=\sum_{i=0}^{\infty}0.5^i x_{t-i}$$

模型 3：$x_t=\varepsilon_t-\frac{4}{5}\varepsilon_{t-1}+\frac{16}{25}\varepsilon_{t-2}$，根据逆函数的递推公式，并根据该模型特有的 $\theta_2=-\theta_1^2$ 的性质，整理之后该模型的逆函数如下：

$$I_k=\begin{cases} (-1)^n\theta_1^k & ,k=3n \text{ 或 } 3n+1 \\ 0 & ,k=3n+2 \end{cases}, n=0,1,\cdots$$

该模型的逆转形式为：

$$\varepsilon_t=\sum_{n=0}^{\infty}(-1)^n 0.8^{3n}x_{t-3n}+\sum_{n=0}^{\infty}(-1)^n 0.8^{3n+1}x_{t-3n-1}$$

3.2.3 *ARMA* 模型

一、定义

定义 3.5 把具有如下结构的模型称为自回归移动平均模型，简记为 $ARMA(p,q)$：

$$\begin{cases} x_t=\phi_0+\phi_1x_{t-1}+\cdots+\phi_px_{t-p}+\varepsilon_t-\theta_1\varepsilon_{t-1}-\cdots-\theta_q\varepsilon_{t-q} \\ \phi_p\neq 0, \theta_q\neq 0 \\ E(\varepsilon_t)=0, \mathrm{Var}(\varepsilon_t)=\sigma_\varepsilon^2, E(\varepsilon_t\varepsilon_s)=0, s\neq t \\ Ex_s\varepsilon_t=0, \forall s<t \end{cases} \tag{3.38}$$

若 $\phi_0=0$，该模型称为中心化 $ARMA(p,q)$模型。缺省默认条件，中心化$ARMA(p,q)$模型可以简写为：

$$x_t=\phi_1x_{t-1}+\cdots+\phi_px_{t-p}+\varepsilon_t-\theta_1\varepsilon_{t-1}-\cdots-\theta_q\varepsilon_{t-q} \tag{3.39}$$

默认条件与 AR 模型、MA 模型相同。

引进延迟算子，$ARMA(p,q)$模型简记为：

$$\Phi(B)x_t=\Theta(B)\varepsilon_t$$

式中：

$\Phi(B)=1-\phi_1B-\cdots-\phi_pB^p$，为 p 阶自回归系数多项式。

$\Theta(B)=1-\theta_1B-\cdots-\theta_qB^q$，为 q 阶移动平均系数多项式。

显然，

当 $q=0$ 时，$ARMA(p,q)$模型就退化成了 $AR(p)$模型；

当 $p=0$ 时，$ARMA(p,q)$模型就退化成了 $MA(q)$模型。

所以，$AR(p)$模型和 $MA(q)$模型实际上是 $ARMA(p,q)$模型的特例，它们都统称为 $ARMA$ 模型。而 $ARMA(p,q)$模型的统计性质也正是 $AR(p)$模型和 $MA(q)$模型统计性质的有机组合。

二、平稳条件与可逆条件

对于一个 $ARMA(p,q)$模型，令 $z_t=\Theta(B)\varepsilon_t$，显然$\{z_t\}$是一个均值为零、方差为$(1+\beta_1^2+\cdots+\beta_q^2)\sigma_\varepsilon^2$ 的平稳序列。于是 $ARMA(p,q)$模型可以改写为如下形式：

$$\Phi(B)x_t=z_t$$

类似于 $AR(p)$模型平稳性的分析，容易推导出 $ARMA(p,q)$模型的平稳条件是：$\Phi(B)=0$ 的根都在单位圆外。也就是说，$ARMA(p,q)$模型的平稳性完全由其自回归部分的平稳性所决定。

同理，可以推导出 $ARMA(p,q)$模型的可逆条件和 $MA(q)$模型的可逆条件完全相同：当 $\Theta(B)=0$ 的根都在单位圆外时，$ARMA(p,q)$模型可逆。

当 $\Phi(B)=0$，$\Theta(B)=0$ 的根都在单位圆外时，$ARMA(p,q)$模型称为平稳可逆模型，这是一个由它的自相关系数惟一识别的模型。

三、传递形式与逆转形式

对于一个平稳可逆 $ARMA(p,q)$模型，它的传递形式为：

$$x_t=\Phi^{-1}(B)\Theta(B)\varepsilon_t=\varepsilon_t+\sum_{j=1}^{\infty}G_j\varepsilon_{t-j}$$

式中，$\{G_1,G_2,\cdots\}$为 Green 函数。

通过待定系数法，容易得到 $ARMA(p,q)$模型场合下 Green 函数的递推公式为：

$$\begin{cases}G_0=1\\ G_k=\sum_{j=1}^{k}\phi'_jG_{k-j}-\theta'_j,k\geqslant 1\end{cases}\tag{3.40}$$

式中：

$$\phi'_j=\begin{cases}\phi_j,1\leqslant j\leqslant p\\ 0,\ j>p\end{cases},\quad \theta'_j=\begin{cases}\theta_j,1\leqslant j\leqslant q\\ 0,\ j>q\end{cases}$$

同理，可以得到 $ARMA(p,q)$模型的逆转形式为：

$$\varepsilon_t=\Theta^{-1}(B)\Phi(B)x_t=x_t+\sum_{j=1}^{\infty}I_jx_{t-j}$$

式中，$\{I_1, I_2, \cdots\}$为逆函数。

通过待定系数法容易得到逆函数的递推公式为：

$$\begin{cases} I_0 = 1 \\ I_k = \sum_{j=1}^{k} \theta'_j I_{k-j} - \phi'_j, k \geq 1 \end{cases} \tag{3.41}$$

式中，θ'_j和ϕ'的定义同上。

四、$ARMA(p,q)$模型的统计性质

1. 均值

对于一个非中心化平稳可逆的$ARMA(p,q)$模型

$$x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

两边同求均值，有

$$Ex_t = \frac{\phi_0}{1 - \phi_1 - \cdots - \phi_p}$$

2. 自协方差函数

$$\begin{aligned} \gamma(k) &= E(x_t x_{t+k}) \\ &= E\Big[\Big(\sum_{i=0}^{\infty} G_i \varepsilon_{t-i}\Big)\Big(\sum_{j=0}^{\infty} G_j \varepsilon_{t+k-j}\Big)\Big] \\ &= E\Big[\sum_{i=0}^{\infty} G_i \sum_{j=0}^{\infty} G_j \varepsilon_{t-i} \varepsilon_{t+k-j}\Big] \\ &= \sigma_\varepsilon^2 \sum_{i=0}^{\infty} G_i G_{i+k} \end{aligned}$$

3. 自相关系数

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)} = \frac{\sum_{j=0}^{\infty} G_j G_{j+k}}{\sum_{j=0}^{\infty} G_j^2}$$

根据自相关系数的表达式很容易判断$ARMA(p,q)$模型的自相关系数不截尾。这和$ARMA(p,q)$模型可以转化为无穷阶移动平均模型的性质是一致的。同理，根据$ARMA(p,q)$模型可以转化为无穷阶自回归模型，可以判断它的偏相关系数也不截尾。

【例 3.7】 拟合$ARMA(1,1)$模型：$x_t - 0.5x_{t-1} = \varepsilon_t - 0.8\varepsilon_t$，并直观地考察该模型自相关系数和偏自相关系数的拖尾性。假定$\{\varepsilon_t\}$为标准正态白噪声序列。

拟合这个$ARMA(1,1)$模型，并得到样本自相关图和偏自相关图如图 3—7 所示。

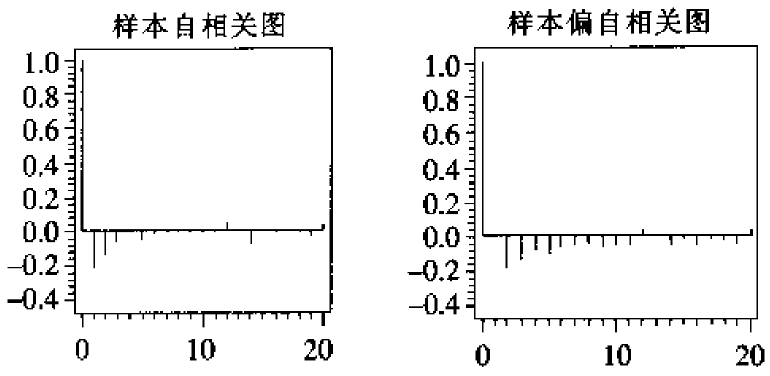


图 3—7 *ARMA*(1,1)模型的样本相关图

综合考察 $AR(p)$模型、$MA(q)$模型和 $ARMA(p,q)$模型自相关系数和偏自相关系数的性质，我们可以总结出如下规律，见表 3—4。

表 3—4

模型	自相关系数	偏自相关系数
$AR(p)$	拖尾	p 阶截尾
$MA(q)$	q 阶截尾	拖尾
$ARMA(p,q)$	拖尾	拖尾

3.3 平稳序列建模

3.3.1 建模步骤

假如某个观察值序列通过序列预处理，可以判定为平稳非白噪声序列，我们就可以利用模型对该序列建模。建模的基本步骤如图 3—8 所示。

1. 求出该观察值序列的样本自相关系数(ACF)和样本偏自相关系数(PACF)的值。

2. 根据样本自相关系数和偏自相关系数的性质，选择阶数适当的$ARMA(p,q)$模型进行拟合。

3. 估计模型中未知参数的值。

4. 检验模型的有效性。如果拟合模型通不过检验，转向步骤 2，重新选择模型再拟合。

5. 模型优化。如果拟合模型通过检验，仍然转向步骤 2，充分考虑各种可能，建立多个拟合模型，从所有通过检验的拟合模型中选择最优模型。

6. 利用拟合模型，预测序列的将来走势。

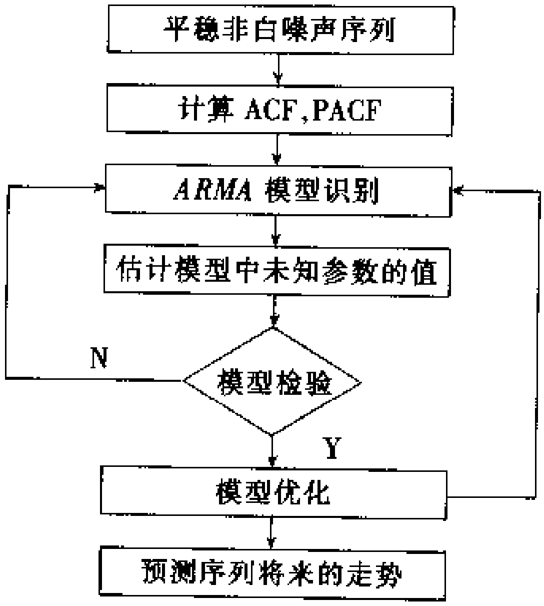


图 3—8　建模步骤

3.3.2　样本自相关系数与偏自相关系数

我们是通过考察平稳序列样本自相关系数和偏自相关系数的性质选择适合的模型拟合观察值序列，所以模型拟合的第一步是要根据观察值序列的取值求出该序列的样本自相关系数$\{\hat{\rho}_k, 0<k<n\}$和样本偏自相关系数$\{\hat{\phi}_{kk}, 0<k<n\}$的值。

样本自相关系数可以根据以下公式求得：

$$\hat{\rho}_k \frac{\sum_{t=1}^{n-k}(x_t-\bar{x})(x_{t+k}-\bar{x})}{\sum_{t=1}^{n}(x_t-\bar{x})^2}, \forall 0<k<n$$

样本偏自相关系数可以利用样本自相关系数的值，根据以下公式求得：

$$\hat{\phi}_{kk}=\frac{\hat{D}_k}{\hat{D}}, \forall 0<k<n$$

式中：

$$\hat{D}=\begin{vmatrix} 1 & \hat{\rho}_1 & \cdots & \hat{\rho}_{k-1} \\ \hat{\rho}_1 & 1 & \cdots & \hat{\rho}_{k-2} \\ \vdots & \vdots & & \vdots \\ \hat{\rho}_{k-1} & \hat{\rho}_{k-2} & \cdots & 1 \end{vmatrix}, \hat{D}=\begin{vmatrix} 1 & \hat{\rho}_1 & \cdots & \hat{\rho}_1 \\ \hat{\rho}_1 & 1 & \cdots & \hat{\rho}_2 \\ \vdots & \vdots & & \vdots \\ \hat{\rho}_{k-1} & \hat{\rho}_{k-2} & \cdots & \hat{\rho}_k \end{vmatrix}$$

3.3.3 模型识别

计算出样本自相关系数和偏自相关系数的值之后，就要根据它们表现出来的性质，选择适当的 *ARMA* 模型拟合观察值序列。这个过程实际上就是要根据样本自相关系数和偏自相关系数的性质估计自相关阶数 $\hat{p}$ 和移动平均阶数 $\hat{q}$，因此，模型识别过程也称为模型定阶过程。

ARMA 模型定阶的基本原则如表 3—5 所示。

表 3—5

$\hat{\rho}_k$	$\hat{\phi}_{kk}$	模型定阶
拖尾	p 阶截尾	$AR(p)$模型
q 阶截尾	拖尾	$MA(q)$模型
拖尾	拖尾	$ARMA(p,q)$模型

但是在实践中，这个定阶原则在操作上具有一定的困难。因为由于样本的随机性，样本的相关系数不会呈现出理论截尾的完美情况，本应截尾的样本自相关系数或偏自相关系数仍会呈现出小值振荡的情况。同时，由于平稳时间序列通常都具有短期相关性，随着延迟阶数 $k\rightarrow\infty$，$\hat{\rho}_k$ 与 $\hat{\phi}_{kk}$ 都会衰减至零值附近作小值波动。

这种现象导致我们必须思考，当样本自相关系数或偏自相关系数在延迟若干阶之后衰减为小值波动时，什么情况下该看做相关系数截尾，什么情况下该看做相关系数在延迟若干阶之后正常衰减到零值附近作拖尾波动呢？

这实际上没有绝对的标准，很大程度上依靠分析人员的主观经验。但样本自相关系数和偏自相关系数的近似分布可以帮助缺乏经验的分析人员做出尽量合理的判断。

Jankins 和 Watts 于 1968 年证明

$$E(\hat{\rho}_k)=\left(1-\frac{k}{n}\right)\rho_k$$

也就是说该样本自相关系数是总体自相关系数的有偏估计值。当 k 足够大时，根据平稳序列自相关系数呈负指数衰减，有 $\rho_k\rightarrow 0$。

根据 Bartlett 公式计算样本自相关系数的方差近似等于：

$$\mathrm{Var}(\hat{\rho}_k)\cong\frac{1}{n}\sum_{m=-j}^{j}\hat{\rho}_m^2=\frac{1}{n}(1+2\sum_{m=-j}^{j}\hat{\rho}_m^2),k>j$$

当样本容量 n 充分大时，样本自相关系数近似服从正态分布：

$$\hat{\rho}_k\sim N(0,\frac{1}{n})$$

Quenouille 证明，样本偏自相关系数也同样近似服从这个正态分布：

$$\hat{\phi}_{kk}\sim N(0,\frac{1}{n})$$

根据正态分布的性质，有

$$\Pr\left(-\frac{2}{\sqrt{n}}\leqslant\hat{\rho}_k\leqslant\frac{2}{\sqrt{n}}\right)\geqslant 0.95$$

$$\Pr\left(-\frac{2}{\sqrt{n}}\leqslant\hat{\phi}_{kk}\leqslant\frac{2}{\sqrt{n}}\right)\geqslant 0.95$$

所以可以利用 2 倍标准差范围辅助判断。

如果样本自相关系数或偏自相关系数在最初的 d 阶明显大于 2 倍标准差范围，而后几乎 95%的自相关系数都落在 2 倍标准差的范围以内，而且由非零自相关系数衰减为小值波动的过程非常突然，这时，通常视为自相关系数截尾。截尾阶数为 d。

如果有超过 5%的样本相关系数落入 2 倍标准差范围之外，或者是由显著非零的相关系数衰减为小值波动的过程比较缓慢或者非常连续，这时，通常视为相关系数不截尾。

【例 2.5 续】 选择合适的 *ARMA* 模型拟合 1950—1998 年北京市城乡居民定期储蓄比例序列。

根据例 2.5 的分析已知该序列为平稳非白噪声序列。可以对该序列拟合 *ARMA* 模型。

考察该序列的样本自相关图（图 2—10），自相关图显示延迟 3 阶之后，自相关系数全部衰减到 2 倍标准差范围内波动，这表明序列明显地短期相关。但序列由显著非零的相关系数衰减为小值波动的过程相当连续、相当缓慢，该自相关系数可视为不截尾。

再考察该序列样本偏自相关系数图，见图 3—9。

```
                         Partial Autocorrelations
Lag  Correlation   -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  1    0.702 46     |             .     |**************     |
  2    0.201 24     |             .     |**** .             |
  3    0.005 12     |             .     |     .             |
  4   -0.126 11     |             .  ***|     .             |
  5    0.226 98     |             .     |*****.             |
  6    0.011 90     |             .     |     .             |
  7    0.030 00     |             .     |*    .             |
  8   -0.262 41     |             .*****|     .             |
  9   -0.062 06     |             .    *|     .             |
 10   -0.104 68     |             .   **|     .             |
 11    0.078 79     |             .     |**   .             |
 12    0.055 19     |             .     |*    .             |
```

图 3—9 北京市城乡居民定期储蓄比例序列的样本偏自相关图

图 3—9 显示除了延迟 1 阶的偏自相关系数显著大于 2 倍标准差之外，其他的偏自相关系数都在 2 倍标准差范围内作小值随机波动，而且由非零相关系数衰减为小值波动的过程非常突然，所以该偏自相关系数可视为 1 阶截尾。

因此，我们可以考虑用 $AR(1)$ 模型：$x_t=\mu+\frac{\epsilon_t}{1-\phi_1 B}$，$\epsilon_t \overset{i.i.d}{\sim} N(0,\sigma_\epsilon^2)$ 来拟合该观察值序列。

【例 3.8】 选择合适的 *ARMA* 模型拟合美国科罗拉多州某一加油站连续 57 天的 OVERSHORT 序列（数据见附录 1.6）。

首先绘制该序列时序图，直观检验序列的平稳性。见图 3—10。

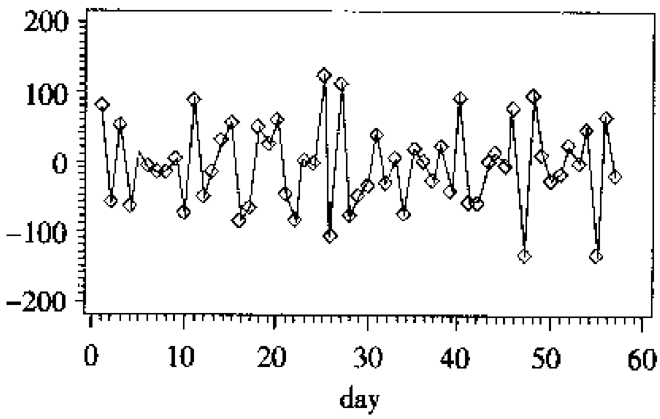


图 3—10 OVERSHORT 序列时序图

时序图显示序列没有显著的非平稳特征。再考察自相关图和偏自相关图，进一步确定平稳性并给拟合模型定阶。见图 3—11 和图 3--12。

```
                                              Autocorrelations
Lag  Covariance   Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
 0    3 415.718    1.000 00     |                    |********************|
 1   -1 719.956    -.503 54     |          **********|   .                |
 2      416.701    0.122 00     |            .       |**  .               |
 3     -723.256    -.211 74     |            .   ****|    .               |
 4      273.522    0.080 08     |            .       |**  .               |
 5    66.440 027   0.019 45     |            .       |    .               |
 6      396.723    0.116 15     |            .       |**  .               |
 7     -742.192    -.217 29     |            .   ****|    .               |
 8      861.426    0.252 19     |            .       |*****.              |
 9     -655.675    -.191 96     |            .   ****|    .               |
10      192.042    0.056 22     |            .       |*   .               |
11     -357.536    -.104 67     |            .     **|    .               |
12    42.072 379   0.012 32     |            .       |    .               |
13      747.289    0.218 78     |            .       |****.               |
14     -203.674    -.059 63     |           .       *|     .              |
```

"·" marks two standard errors

图 3—11 OVERSHORT 序列自相关图

```
                         Partial Autocorrelations
Lag  Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  1  -0.503 54     |          **********|    .               |
  2  -0.176 25     |               .****|    .               |
  3  -0.315 56     |              ******|    .               |
  4  -0.263 67     |               *****|    .               |
  5  -0.152 82     |               . ***|    .               |
  6   0.042 00     |               .    |*   .               |
  7  -0.193 96     |               .****|    .               |
  8   0.118 50     |               .    |**  .               |
  9   0.050 68     |               .    |*   .               |
 10  -0.069 69     |               .   *|    .               |
 11  -0.152 94     |               . ***|    .               |
 12  -0.248 58     |               *****|    .               |
 13   0.097 86     |               .    |**  .               |
 14   0.102 70     |               .    |**  .               |
```

图 3—12 OVERSHORTS 序列偏自相关图

自相关图显示除了延迟 1 阶的自相关系数在 2 倍标准差范围之外，其他阶数的自相关系数都在 2 倍标准差范围内波动。根据自相关系数的这个特点可以判断该序列具有短期相关性，进一步确定序列平稳。同时，可以认为该序列自相关系数 1 阶截尾。

偏自相关系数显示出非截尾的性质。综合该序列自相关系数和偏自相关系数的性质，为拟合模型定阶为 $MA(1)$。

【例 3.9】 选择合适的 $ARMA$ 模型拟合 1880—1985 年全球气表平均温度改变值差分序列（全球气表平均温度改变值序列见附录 1.22）。

对 1880—1985 年全球气表平均温度改变值序列进行差分运算，并对差分后序列绘制时序图，见图 3—13。

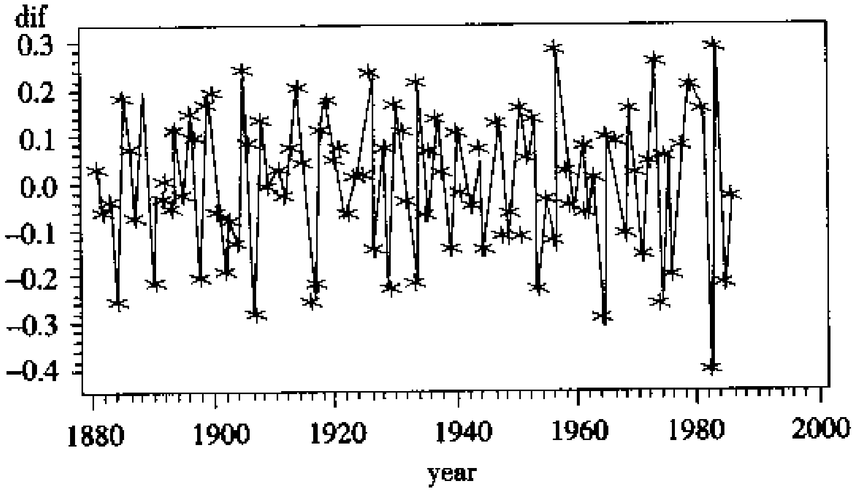


图 3—13 全球气表平均温度改变偏差分序列时序图

时序图显示序列的波动非常平稳，再考察其自相关图及偏自相关图的性质，见图 3 -14 和图 3—15。

```
                                  Autocorrelations
Lag  Covariance     Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
 0    0.020 061     1.000 00      |                    |********************|
 1   -0.005 073 2   -.252 89      |               *****|   .                |
 2   -0.002 277 8   -.113 54      |                . **|   .                |
 3   -0.003 178 6   -.158 45      |                .***|   .                |
 4    0.000 893 72  0.044 55      |                .   |*  .                |
 5   -0.001 993 0   -.099 35      |                . **|   .                |
 6    0.004 317 7   0.215 23      |                .   |****                |
 7   -0.004 056 7   -.202 22      |                ****|   .                |
 8    0.002 104 5   0.104 90      |               .    |**  .               |
 9   -0.002 119 5   -.105 66      |               .  **|    .               |
10    0.002 066 4   0.103 01      |               .    |**  .               |
11   -0.002 314     -.011 54      |               .    |    .               |
12    0.000 867 29  0.043 23      |               .    |*   .               |
                          "." marks two standard errors
```

图 3—14　全球气表平均温度改变值差分序列自相关图

```
                        Partial Autocorrelations
Lag  Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
 1    -0.252 89    |               *****|   .                |
 2    -0.189 63    |                ****|   .                |
 3    -0.266 59    |               *****|   .                |
 4    -0.130 25    |                .***|   .                |
 5    -0.241 34    |               *****|   .                |
 6    -0.055 58    |                .   |*  .                |
 7    -0.230 49    |               *****|   .                |
 8    -0.031 22    |                .  *|   .                |
 9    -0.159 70    |                .***|   .                |
10    -0.048 92    |                .  *|   .                |
11    -0.014 93    |                .   |   .                |
12    -0.053 20    |                .  *|   .                |
```

图 3—15　全球气表平均温度改变值差分序列偏自相关图

自相关系数和偏自相关系数均显示出不截尾的性质，因此可以尝试使用 $ARMA(1,1)$ 拟合该序列。

3.3.4　参数估计

选择好了拟合模型之后，下一步就是要利用序列的观察值确定该模型的口径，即估计模型中未知参数的值。

对于一个非中心化 $ARMA(p,q)$ 模型，有

$$x_t=\mu+\frac{\Theta_q(B)}{\Phi_p(B)}\varepsilon_t$$

式中：

$$\varepsilon_t \sim WN(0,\sigma_\varepsilon^2)$$

$$\Phi_p(B)=1-\theta_1 B-\cdots-\theta_p B^p$$

$$\Theta_q(B)=1-\phi_1 B-\cdots-\phi_q B^q$$

该模型共含有 $p+q+2$ 个未知参数：$\phi_1,\cdots,\phi_p,\theta_1,\cdots,\theta_q,\mu,\sigma_\varepsilon^2$。

参数 μ 是序列均值，通常采用矩估计方法，用样本均值估计总体均值即可得到它的估计值：

$$\hat{\mu}=\bar{x}=\frac{\sum_{i=1}^{n} x_i}{n}$$

对原序列中心化，有

$$y_t=x_t-\bar{x}$$

原 $p+q+2$ 个待估参数减少为 $p+q+1$ 个：ϕ_1，…，ϕ_p，θ_1，…，θ_q，σ_ε^2。对这 $p+q+1$ 个未知参数的估计方法有三种：矩估计、极大似然估计和最小二乘估计。

一、矩估计

运用 $p+q$ 个样本自相关系数估计总体自相关系数

$$\begin{cases}\rho_1(\phi_1,\cdots,\phi_p,\theta_1,\cdots,\theta_q)=\hat{\rho}_1\\ \cdots\cdots\cdots\cdots\\ \rho_{p+q}(\phi_1,\cdots,\phi_p,\theta_1,\cdots,\theta_q)=\hat{\rho}_{p+q}\end{cases}$$

从中解出的参数值 $\hat{\phi}_1,\cdots,\hat{\phi}_p$，$\hat{\theta}_1,\cdots,\hat{\theta}_q$ 就是 $\phi_1,\cdots,\phi_p,\theta_1,\cdots,\theta_q$ 的矩估计。

用序列样本方差估计序列总体方差

$$\hat{\sigma}_x^2=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$$

在 $ARMA(p,q)$ 模型两边同时求方差，整理得到 σ_ε^2 关于 σ_x^2 的函数形式

$$\sigma_\varepsilon^2=\frac{1+\phi_1^2+\cdots+\phi_p^2}{1+\theta_1^2+\cdots+\theta_q^2}\sigma_x^2 \tag{3.42}$$

把 $\hat{\phi}_1,\cdots,\hat{\phi}_p$，$\hat{\theta}_1,\cdots,\hat{\theta}_q$ 及 $\hat{\sigma}_x^2$ 代入式(3.42)，即得到白噪声序列方差的矩估计

$$\hat{\sigma}_\varepsilon^2=\frac{1+\hat{\phi}_1^2+\cdots+\hat{\phi}_p^2}{1+\hat{\theta}_1^2+\cdots+\hat{\theta}_q^2}\hat{\sigma}_x^2$$

【例 3.10】 求 $AR(2)$ 模型 $x_t=\phi_1 x_{t-1}+\phi_2 x_{t-2}+\varepsilon_t$ 中未知参数 ϕ_1，ϕ_2 的矩估计。

根据 Yule Walker 方程，有

$$\begin{cases}\rho_1=\phi_1+\phi_2\rho_1\\ \rho_2=\phi_1\rho_1+\phi_2\end{cases}$$

则

$$\hat{\phi}_1=\frac{1-\hat{\rho}_2}{1-\hat{\rho}_1}\hat{\rho}_1,\hat{\phi}_2=\frac{\hat{\rho}_2-\hat{\rho}_1{}^2}{1-\hat{\rho}_1{}^2}$$

【例 3.11】 求 $MA(1)$模型 $x_t=\varepsilon_t-\theta_1\varepsilon_{t-1}$中未知参数 θ_1 的矩估计。

根据 $MA(1)$模型自协方差函数的性质，有

$$\begin{cases}\gamma_0=(1+\theta_1^2)\sigma_\varepsilon^2\\ \gamma_1=-\theta_1\sigma_\varepsilon^2\end{cases}\Rightarrow\rho_1=\frac{\gamma_1}{\gamma_0}=\frac{-\theta_1}{1+\theta_1^2}$$

解一元二次方程

$$\theta_1^2\rho_1+\theta_1+\rho_1=0$$

得

$$\hat{\theta}_1=\frac{-1\pm\sqrt{1-4\hat{\rho}_1^2}}{2\hat{\rho}_1}$$

考虑 $MA(1)$模型的可逆性条件：$|\theta_1|<1$，可得到未知参数的惟一解：

$$\hat{\theta}_1=\frac{-1+\sqrt{1-4\hat{\rho}_1^2}}{2\hat{\rho}_1}$$

【例 3.12】 求 $ARMA(1,1)$模型 $x_t=\phi_1x_{t-1}+\varepsilon_t-\theta_1\varepsilon_{t-1}$中未知参数 ϕ_1，θ_1 的矩估计。

根据 $ARMA$ 模型 Green 函数的递据公式，可以确定该 $ARMA(1,1)$模型的 Green 函数为：

$$G_0=1$$

$$G_k=(\phi_1-\theta_1)\ \phi_1^{k-1},\ k=1,\ 2,\ \cdots$$

推导出

$$\begin{cases}\gamma_0=\sum_{k=0}^{\infty}G_k^2\sigma_\varepsilon^2=\frac{1+\theta_1^2-2\theta_1\phi_1}{1-\phi_1^2}\sigma_\varepsilon^2\\ \gamma_1=\sum_{k=0}^{\infty}G_kG_{k+1}\sigma_\varepsilon^2=\frac{(\phi_1-\theta_1)(1-\theta_1\phi_1)}{1-\phi_1^2}\sigma_\varepsilon^2\\ \gamma_2=\sum_{k=0}^{\infty}G_KG_{k+2}\sigma_\varepsilon^2=\phi_1\gamma_1\end{cases}$$

则

$$\begin{cases}\rho_1=\frac{\gamma_1}{\gamma_0}=\frac{(\phi_1-\theta_1)(1-\theta_1\phi_1)}{1+\theta_1^2-2\theta_1\phi_1}\\ \rho_2=\phi_1\rho_1\end{cases}\tag{3.43}$$

整理方程组（3.43），得

$$\begin{cases}\theta_1^2-\dfrac{1+\phi_1^2-2\rho_2}{\phi_1-\rho_1}\theta_1+1=0\\ \phi_1=\dfrac{\rho_2}{\rho_1}\end{cases}$$

考虑可逆条件：$|\theta_1|<1$，得到未知参数矩估计的惟一解

$$\begin{cases}\hat{\phi}_1=\dfrac{\hat{\rho}_2}{\hat{\rho}_1}\\ \hat{\theta}_1=\begin{cases}\dfrac{c+\sqrt{c^2-4}}{2}, & c\leqslant-2\\ \dfrac{c-\sqrt{c^2-4}}{2}, & c\geqslant 2\end{cases}, \quad c=\dfrac{1-\phi_1^2-2\hat{\rho}_2}{\phi_1-\hat{\rho}_1}\end{cases}$$

矩估计方法，尤其是低阶 *ARMA* 模型场合下的矩估计方法具有计算量小、估计思想简单直观，且不需要假设总体分布的优点。但是在这种估计方法中只用到了 $p+q$ 个样本自相关系数，即样本二阶矩的信息，观察值序列中的其他信息都被忽略了。这导致矩估计方法是一种比较粗糙的估计方法，它的估计精度一般较差，因此它常被用做极大似然估计和最小二乘估计迭代计算的初始值。

二、极大似然估计

在极大似然准则下，认为样本来自使该样本出现概率最大的总体。因此未知参数的极大似然估计就是值得似然函数（即联合密度函数）达到最大的参数值。

$$L(\hat{\beta}_1,\hat{\beta}_2,\cdots,\hat{\beta}_k;x_1,x_2,\cdots,x_n)=\max\{p(x_1,x_2,\cdots,x_n);\beta_1,\beta_2,\cdots,\beta_k\} \quad (3.44)$$

使用极大似然估计必须已知总体的分布函数，而在时间序列分析中，序列总体的分布通常是未知的。为了值于分析和计算，通常假设序列服从多元正态分布。

$$x_t=\phi_1x_{t-1}+\cdots+\phi_px_{t-p}+\varepsilon_t-\theta_1\varepsilon_{t-1}-\cdots-\theta_q\varepsilon_{t-q}$$

记

$$\tilde{x}=(x_1,\cdots,x_n)'$$

$$\tilde{\beta}=(\phi_1,\cdots,\phi_p,\theta_1,\cdots,\theta_q)'$$

$$\sum\nolimits_n=E(\tilde{x}'\tilde{x})=\Omega\sigma_\varepsilon^2$$

式中：

$$\Omega=\begin{pmatrix}\sum\limits_{i=0}^{\infty}G_i^2 & \cdots & \sum\limits_{i=0}^{\infty}G_iG_{i+n-i}\\ \vdots & \ddots & \vdots\\ \sum\limits_{i=0}^{\infty}G_iG_{i+n-1} & \cdots & \sum\limits_{i=0}^{\infty}G_i^2\end{pmatrix}$$

$\tilde{x}$ 的似然函数为：

$$L(\tilde{\beta};\tilde{x}) = p(x_1,x_2,\cdots,x_n;\tilde{\beta})$$

$$= (2\pi)^{-\frac{n}{2}} \mid \sum_n \mid^{\frac{1}{2}} \exp\{-\frac{\tilde{x}\sum_n^{-1}\tilde{x}}{2}\}$$

$$= (2\pi)^{-\frac{n}{2}}(\sigma_\varepsilon^2)^{-\frac{n}{2}} \mid \Omega \mid^{-\frac{1}{2}} \exp\{-\frac{\tilde{x}'\Omega\tilde{x}}{2\sigma_\varepsilon^2}\}$$

对数似然函数为

$$l(\tilde{\beta};\tilde{x}) = -\frac{n}{2}\ln(2\pi) - -\frac{n}{2}\ln(\sigma_\varepsilon^2) - \frac{1}{2}\ln|\Omega| - \frac{1}{2\sigma_\varepsilon^2}[\tilde{x}'\Omega^{-1}\tilde{x}]$$

对对数似然函数中得未知参数求偏导数，得到似然方程组

$$\begin{cases} \frac{\partial}{\partial\sigma_\varepsilon^2} l(\tilde{\beta};\tilde{x}) = \frac{n}{2\sigma_\varepsilon^2} - \frac{S(\tilde{\beta})}{2\sigma_\varepsilon^4} = 0 \\ \frac{\partial}{\partial\tilde{\beta}} l(\tilde{\beta};\tilde{x}) = \frac{1}{2}\frac{\partial\ln|\Omega|}{\partial\tilde{\beta}} + \frac{1}{\sigma_\varepsilon^2}\frac{\partial S(\tilde{\beta})}{2\partial\tilde{\beta}} = 0 \end{cases} \tag{3.45}$$

式中 $S(\tilde{\beta}) = \tilde{x}'\Omega^{-1}\tilde{x}$

理论上，求解方程组（3.45）即得到未知参数的极大似然估计值。但是，由于 $S(\tilde{\beta})$ 和 $\ln|\Omega|$ 都不是 $\tilde{\beta}$ 的显式表达式，因而似然方程组（3.45）实际上是由 $p+q+1$ 个超越方程构成，通常需要经过复杂的迭代算法才能求出未知参数的极大似然估计值。

幸运的是，目前计算机技术比较发达，有许多统计软件可以辅助分析，使得求 $ARMA$ 模型的极大似然估计值成为一件很容易实现的事情。

极大似然估计充分应用了每一个观察值所提供的信息，因而它的估计精度高，同时还具有估计的一致性、渐近正态性和渐近有效性等许多优良的统计性质，是一种非常优良的参数估计方法。

三、最小二乘估计

在 $ARMA(p,q)$ 模型场合，记

$$\tilde{\beta} = (\phi_1,\cdots,\phi_p,\theta_1,\cdots,\theta_q)'$$

$$F_t(\tilde{\beta}) = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}$$

残差项为：

$$\varepsilon_t = x_t - F_t(\tilde{\beta})$$

残差平方和为：

$$Q(\tilde{\beta}) = \sum_{t=1}^{n}\varepsilon_t^2$$

$$= \sum_{t=1}^{n}(x_t - \phi_1 x_{t-1} - \cdots - \phi_p x_{t-p} - \theta_q \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q})$$

使残差平方和达到最小的那组参数值即为 $\tilde{\beta}$ 的最小二乘估计值。

同极大似然估计一样，由于 $Q(\tilde{\beta})$不是 $\tilde{\beta}$ 的显性函数，未知参数的最小二乘估计值通常也得借助迭代法求出。由于充分利用了序列观察值的信息，因而最小二乘估计的精度很高。

在实际运用中，最常用的是条件最小二乘估计方法。它值定过去未观测到的序列值等于零，即

$$x_t = 0, t \leqslant 0$$

根据这个假定可以得到残差序列的有限项表达式：

$$\varepsilon_t = \frac{\Phi(B)}{\Theta(B)} x_t = x_t - \sum_{i=1}^{t} \pi_i x_{t-i}$$

于是残差平方和为：

$$Q(\tilde{\beta}) = \sum_{t=1}^{n} \varepsilon_t^2$$

$$= \sum_{t=1}^{n} \left[x_t - \sum_{i=1}^{t} \pi_i x_{t-1} \right]^2 \tag{3.46}$$

通过迭代法，使式（3.46）达到最小值的估计值即为参数 β 的条件最小二乘估计。

【例 2.5 续】 确定 1950—1998 年北京市城乡居民定期储蓄比例序列拟合模型的口径。

根据该序列自相关图和偏自相关图的性质，我们为该序列定阶为 $AR(1)$模型,使用极大似然估计确定该模型口径为：

$$x_t = 25.17 + 0.69 x_{t-1} + \varepsilon_t, \mathrm{Var}(\hat{\sigma}_\varepsilon^2) = 16.17$$

【例 3.8 续】 确定美国科罗拉多州某一加油站连续 57 天的 OVERSHORTS 序列拟合模型的口径。

在例 3.8 中，我们为该序列的拟合模型定阶为 $MA(1)$模型，使用条件最小二乘估计，容易确定该模型口径为：

$$x_t = -4.403\,51 + (1 - 0.823\,03B)\varepsilon_t, \mathrm{Var}(\hat{\sigma}_\varepsilon^2) = 2\,178.929$$

【例 3.9 续】 确定 1880—1985 年全球气表平均温度改变值差分序列拟合模型的口径。

对序列尝试拟合 $ARMA(1,1)$模型，使用条件最小二乘估计，得到该模型的口径为：

$$x_t = 0.003 + 0.407x_{t-1} + \varepsilon_t - 0.9\varepsilon_{t-1}, \mathrm{Var}(\hat{\sigma}_\varepsilon^2) = 0.016$$

3.3.5 模型检验

确定了拟合模型的口径之后，我们还要对该拟合模型进行必要的检验。

一、模型的显著性检验

模型的显著性检验主要是检验模型的有效性。一个模型是否显著有效主要看它提取的信息是否充分。一个好的拟合模型应该能够提取观察值序列中几乎所有的样本相关信息，换言之，拟合残差项中将不再蕴含任何相关信息，即残差序列应该为白噪声序列。这样的模型称之为显著有效模型。

反之，如果残差序列为非白噪声序列，那就意味着残差序列中还残留着相关信息未被提取，这就说明拟合模型不够有效，通常需要选择其他模型，重新拟合。

所以模型的显著性检验即为残差序列的白噪声检验。原假设和备择假设分别为：

$H_0: \rho_1 = \rho_2 = \cdots = \rho_m = 0, \forall m \geq 1$

H_1：至少存在某个 $\rho_k \neq 0, \forall m \geq 1, k \leq m$

检验统计量为 LB（Ljung-Box）检验统计量：

$$LB = n(n+2)\sum_{k=1}^{m}\left(\frac{\hat{\rho}_k^2}{n-k}\right) \sim \chi^2(m), \forall m > 0$$

如果拒绝原假设，就说明残差序列中还残留着相关信息，拟合模型不显著。如果不能拒绝原假设，就认为拟合模型显著有效。

【例 2.5 续】 检验 1950—1998 年北京市城乡居民定期储蓄比例序列拟合模型的显著性。

残差序列白噪声检验结果如表 3—6 所示。

表 3—6

延迟阶数	LB统计量的值	P值
6	5.83	0.322 9
12	10.28	0.505 0
18	11.38	0.836 1

由于各阶延迟下 LB统计量的 P 值都显著大于 0.05，可以认为这个拟合模型的残差序列属于白噪声序列，即该拟合模型显著有效。

二、参数的显著性检验

参数的显著性检验就是要检验每一个未知参数是否显著非零。这个检验的目

的是为了使模型最精简。

如果某个参数不显著，即表示该参数所对应的那个自变量对因变量的影响不明显，该自变量就可以从拟合模型中删除。最终模型将由一系列参数显著非零的自变量表示。

检验假设：

$$H_0:\beta_j=0\leftrightarrow H_1:\beta_j\neq 0,\forall 1\leqslant j\leqslant m$$

$$E(\hat{\beta})=E[(X'X)^{-1}X'\tilde{y}]=(X'X)^{-1}X'X\beta=\tilde{\beta}$$

$$\begin{aligned}\mathrm{Var}(\hat{\beta})&=\mathrm{Var}[(X'X)^{-1}X'\tilde{y}]=(X'X)^{-1}X'X(X'X)^{-1}\sigma_\varepsilon^2\\&=(X'X)^{-1}\sigma_\varepsilon^2\end{aligned}$$

对于线性拟合模型，记$\hat{\beta}$为$\tilde{\beta}$的最小二乘估计，有

$$\Omega=(X'X)^{-1}=\begin{bmatrix}a_{11} & \cdots & a_{1m}\\ \vdots & \ddots & \vdots\\ a_{m1} & \cdots & a_{mm}\end{bmatrix}$$

在正态分布假定下，第j个未知参数的最小二乘估计使$\hat{\beta}_j$服从正态分布：

$$\hat{\beta}_j\sim N(\beta_j,a_{jj}\sigma_\varepsilon^2),i\leqslant j\leqslant m \tag{3.47}$$

由于σ_ε^2不可观测，用最小残差平方和估计σ_ε^2：

$$\hat{\sigma}_\varepsilon^2=\frac{Q(\tilde{\beta})}{n-m}$$

根据正态分布的性质，有

$$\frac{Q(\tilde{\beta})}{\sigma_\varepsilon^2}\sim\chi^2(n-m) \tag{3.48}$$

由式（3.47）和式（3.48）可以构造出用于检验未知参数显著性的t检验统计量

$$T=\sqrt{n-m}\frac{\hat{\beta}_j-\beta_j}{\sqrt{a_{jj}Q(\tilde{\beta})}}\sim t(n-m)$$

当该检验统计量的绝对值大于自由度为$n-m$的t分布的$1-\frac{\alpha}{2}$分位点，即

$$|T|\geqslant t_{1-\alpha/2}(n-m)$$

或者该检验统计量的P值小于$\frac{\alpha}{2}$或大于$1-\frac{\alpha}{2}$时，拒绝原假设，认为该参数显著。否则，认为该参数不显著。这时，应该剔除不显著参数所对应的自变量重新拟合模型，构造出新的、结构更精练的拟合模型。

【例2.5续】 检验1950—1998年北京市城乡居民定期储蓄比例序列极大似然估计模型的参数是否显著（$\alpha=0.05$）。

检验结果如表3—7所示。

表 3—7

估计方法	均值的检验		ϕ_1 的检验		结论
	t 统计量值	P 值	t 统计量值	P 值	
极大似然	46.12	<0.000 1	6.72	<0.000 1	两参数检验均显著

【例 3.8 续】 对 OVERSHORTS 序列的拟合模型进行检验。

残差序列白噪声检验结果如表 3—8 所示。

表 3—8

延迟阶数	LB 统计量的值	P 值	结论
6	3.15	0.677 2	拟合模型
12	9.05	0.617 1	显著有效

残差检验结果显示残差序列可视为白噪声检验，这说明拟合模型显著有效。

参数检验结果如表 3—9 所示。

表 3—9

估计方法	均值的检验		θ_1 的检验		结论
	t 统计量值	P 值	t 统计量值	P 值	
条件最小二乘	−3.75	<0.000 4	10.60	<0.000 1	两参数检验均显著

【例 3.9 续】 对 1880—1985 年全球气表平均温度改变值差分序列拟合模型进行检验。

残差序列白噪声检验结果如表 3—10 所示。

表 3—10

延迟阶数	LB 统计量的值	P 值	结论
6	5.28	0.259 5	拟合模型显著有效
12	10.30	0.414 7	
18	13.34	0.647 7	

参数检验结果如表 3—11 所示。

表 3—11

估计方法	θ_1 的检验		ϕ_1 的检验		结论
	t 统计量值	P 值	t 统计量值	P 值	
条件最小二乘	16.34	0.000 1	3.5	0.000 7	两参数检验均显著

3.3.6 模型优化

一、问题的提出

当一个拟合模型通过了检验，说明在一定的置信水平下，该模型能有效地拟

合观察值序列的波动，但这种有效模型并不是惟一的。

【例 3.13】 等时间间隔，连续读取 70 个某次化学反应的过程数据，构成一时间序列（数据见附录 1.8）。试对该序列进行模型拟合。（α=0.05）

1. 序列预处理

序列时序图显示此次化学反应过程无明显趋势或周期，波动稳定。见图 3—16。

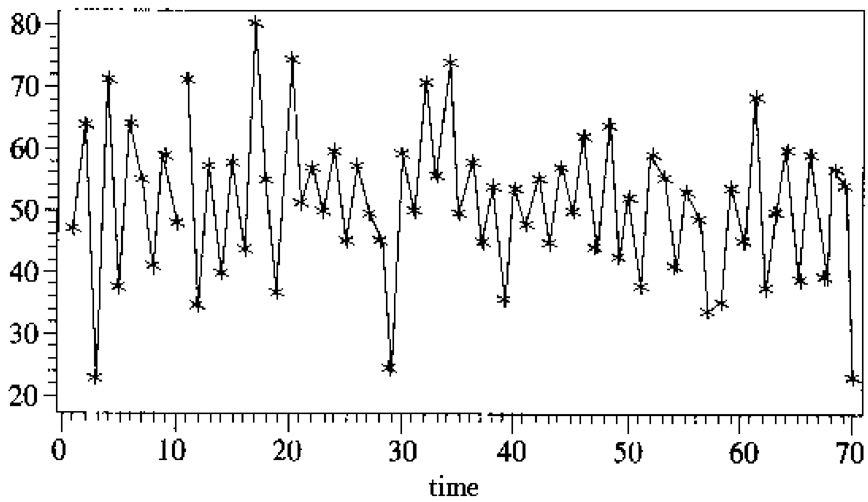


图 3—16 化学反应过程时序图

自相关图显示出自相关系数具有明显的短期相关，2 阶截尾性。见图 3—17。

```
                              Autocorrelations
Lag  Covariance   Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
 0    139.798      1.000 00     |                    |********************|
 1   -54.504 114   -.389 88     |            ********|    .               |
 2    42.553 609   0.304 39     |              .     |******              |
 3   -23.144 178   -.165 55     |             .   ***|     .              |
 4     9.886 402   0.070 72     |             .      |*    .              |
 5   -13.565 875   -0.970 4     |             .    **|     .              |
 6    -6.578 560   -.047 06     |             .     *|     .              |
 7     4.945 082   0.035 37     |             .      |*    .              |
 8    -6.075 359   -.043 46     |             .     *|     .              |
 9    -0.670 493   -.004 80     |             .      |     .              |
10     2.012 128   0.014 39     |             .      |     .              |
11    15.366 178   0.109 92     |             .      |**   .              |
12    -9.615 079   -.068 78     |             .     *|     .              |
13    20.694 889   0.148 03     |             .      |***  .              |
14     5.000 367   0.035 77     |             .      |*    .              |
15    -0.933 542   -.006 68     |             .      |     .              |
16    24.185 609   0.173 00     |             .      |***  .              |
17   -15.565 443   -.111 34     |             .    **|     .              |
18     2.791 872   0.019 97     |             .      |     .              |

                      "." marks two standard errors
```

图 3—17 化学反应过程自相关图

序列随机性检验显示该序列为非白噪声序列。

Autocorrelation Check for White Noise

To Lag	Chi-Square	DF	Pr>ChiSq	Autocorrelations					
6	21.32	6	0.001 6	−0.390	0.304	−0.166	0.071	−0.097	−0.047
12	23.03	12	0.027 4	0.035	−0.043	−0.005	0.014	0.110	−0.069
18	29.10	18	0.047 1	0.148	0.036	−0.007	0.173	−0.111	0.020

综合序列时序图、自相关图和白噪声检验结果，判定该序列为平稳非白噪声序列。可以考虑使用 *ARMA* 模型对它进行拟合。

2. 模型定阶

根据自相关图显示的自相关系数的2阶截尾的性质，尝试拟合 *MA*(2)模型。

3. 参数估计

使用条件最小二乘估计方法，确定 *MA*(2)模型的口径为：

$$yield_t = 51.173\,01 + (1 + 0.322\,86B + 0.310\,09B^2)\varepsilon_t$$

其中 $\sigma_\varepsilon^2 = 119.565\,3$

4. 模型检验

残差白噪声检验显示延迟6阶、12阶、18阶LB检验统计量的 *P* 值均显著大于0.05，所以该 *MA*(2)模型显著有效。

Autocorrelation Check for of Residuals

To Lag	Chi-Square	DF	Pr>ChiSq	Autocorrelations					
6	2.28	4	0.684 2	−0.034	0.024	−0.129	−0.011	−0.086	−0.064
12	4.46	10	0.924 2	0.047	−0.059	−0.046	0.086	0.098	−0.032
18	10.79	16	0.822 5	0.090	0.057	0.092	0.181	−0.101	−0.073
24	13.71	22	0.911 5	−0.043	0.035	0.048	−0.075	−0.062	−0.112

参数显著性检验结果显示三参数 *t* 统计量的 *P* 值均小于0.05，即三参数均显著。

Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr>|t|	Lag
MU	51.173 01	1.284 30	39.84	<.000 1	0
MA1，1	0.322 86	0.121 55	2.66	0.009 9	1
MA1，2	−0.310 09	0.122 04	−2.54	0.013 4	2

因此 *MA*(2)模型是该序列的有效拟合模型。

再返回模型定阶阶段，考察该序列偏自相关系数的性质。偏自相关图显示该

序列偏自相关系数 1 阶截尾。见图 3—18。

```
                        Paartial Autocorrelations

Lag  Correlation   -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  1   -0.389 88    |            ********|    .               |
  2    0.179 71    |               .    |****.               |
  3    0.002 26    |               .    |    .               |
  4   -0.044 28    |               .   *|    .               |
  5   -0.069 41    |               .   *|    .               |
  6   -0.120 62    |               .  **|    .               |
  7    0.019 68    |               .    |    .               |
  8    0.004 89    |               .    |    .               |
  9   -0.056 50    |               .   *|    .               |
 10    0.003 71    |               .    |    .               |
 11    0.142 80    |               .    |*** .               |
 12   -0.009 41    |               .    |    .               |
 13    0.091 96    |               .    |**  .               |
 14    0.166 93    |               .    |*** .               |
 15   -0.001 29    |               .    |    .               |
 16    0.220 69    |               .    |****.               |
 17    0.052 81    |               .    |*   .               |
 18   -0.105 19    |               .  **|    .               |
```

图 3—18 化学反应过程偏自相关图

根据偏自相关系数 1 阶截尾性，尝试拟合 $AR(1)$模型。使用条件最小二乘估计方法确定模型的口径为：

$$yield_t = 51.261\,69 + \frac{\varepsilon_t}{1 + 0.424\,81B}$$

且 $\sigma_\varepsilon^2 = 120.073\,5$

残差白噪声检验显示拟合模型有效。

Autocorrelation Check for of Residuals

To Lag	Chi-Square	DF	Pr>ChiSq	Autocorrelations					
6	4.60	5	0.467 0	0.072	0.155	−0.043	−0.046	−0.122	−0.111
12	7.00	11	0.799 1	0.015	−0.057	−0.031	0.094	0.120	0.027
18	14.45	17	0.634 7	0.157	0.094	0.093	0.179	−0.073	−0.043
24	18.24	23	0.744 3	−0.047	−0.010	−0.005	−0.110	−0.077	−0.123

参数显著性检验结果显示两个参数均显著。

Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr>|t|	Lag
MU	51.261 69	0.922 86	55.55	<.000 1	0
MR1，1	−0.424 81	0.115 61	−3.67	0.000 5	1

因此 $AR(1)$模型也是该序列的有效拟合模型。

同一个序列可以构造两个拟合模型，两个模型都显著有效，那么到底该选择哪个模型用于统计推断呢?

为了解决这个问题，引进 AIC 和 SBC 信息准则的概念，进行模型优化。

二、AIC 准则

AIC 准则是由日本统计学家赤池弘次（Akaike）于 1973 年提出的，它的全称是最小信息量准则（an information criterion）。

该准则的指导思想是认为一个拟合模型的好坏可以从两方面去考察：一方面是大家非常熟悉的、常用来衡量拟合程度的似然函数值；另一方面是模型中未知参数的个数。

通常似然函数值越大说明模型拟合的效果越好。模型中未知参数个数越多，说明模型中包含的自变量越多，自变量越多，模型变化越灵活，模型拟合的准确度就会越高。模型拟合程度高是我们所希望的，但是我们又不能单纯地以拟合精度来衡量模型的好坏，因为这样势必会导致未知参数的个数越多越好。

未知参数越多，说明模型中自变量越多，未知的风险越多。而且参数越多，参数估计的难度就越大，估计的精度也越差。所以一个好的拟合模型应该是一个拟合精度和未知参数个数的综合最优配置。

AIC 准则就是在这种考虑下提出的，它是拟合精度和参数个数的加权函数：

$$\text{AIC}=-2\ln(\text{模型的极大似然函数值})+2(\text{模型中未知参数个数})$$

使 AIC 函数达到最小的模型被认为是最优模型。

在 $ARMA(p,q)$模型场合，对数似然函数为：

$$l(\tilde{\beta};x_1,\cdots,x_n)=-\left[\frac{n}{2}\ln(\sigma_\varepsilon^2)+\frac{1}{2}\ln|\Omega|+\frac{1}{2\sigma_\varepsilon^2}S(\tilde{\beta})\right]$$

因为$\frac{1}{2}\ln|\Omega|$有界，$\frac{1}{2\sigma_\varepsilon^2}S(\tilde{\beta})\to\frac{n}{2}$，所以

$$l(\tilde{\beta};x_1,\cdots,x_n)=\infty-\frac{n}{2}\ln(\sigma_\varepsilon^2)$$

中心化 $ARMA(p,q)$模型的未知参数个数为 $p+q+1$，非中心化 $ARMA(p,q)$模型的未知参数个数为 $p+q+2$。

所以，中心化 $ARMA(p,q)$模型的 AIC 函数为：

$$\text{AIC}=n\ln(\hat{\sigma}_\varepsilon^2)+2(p+q+1)$$

非中心化 $ARMA(p,q)$模型的 AIC 函数为：

$$\text{AIC}=n\ln(\hat{\sigma}_\varepsilon^2)+2(p+q+2)$$

三、SBC 准则

AIC 准则为选择最优模型带来了很大的方便，但 AIC 准则也有不足之处。对于一个观察值序列而言，序列越长，相关信息就越分散，要很充分地提取其中的有用信息，或者说要使得拟合精度比较高的话，通常需要多自变量复杂模型。在 AIC 准则中拟合误差提供的信息要受到样本容量的放大，它等于 $n\ln(\hat{\sigma}_\varepsilon^2)$，但参数个数的惩罚因子却和样本容量没关系，它的权重始终是常数 2。因此在样本容量趋于无穷大时，由 AIC 准则选择的模型不收敛于真实模型，它通常比真实模型所含的未知参数个数要多。

为了弥补 AIC 准则的不足，Akaike 于 1976 年提出 BIC 准则。而 Schwartz 在 1978 年根据 Bayes 理论也得出同样的判别准则，称为 SBC 准则。SBC 准则定义为：

$$\text{SBC}=-2\ln(\text{模型的极大似然函数值})+\ln(n)(\text{模型中未知参数个数})$$

它对 AIC 的改进就是将未知参数个数的惩罚权重由常数 2 变成了样本容量的对数函数 $\ln(n)$。理论上已证明，SBC 准则是最优模型的真实阶数的相合估计。

容易得到，中心化 $ARMA$ (p，q) 模型的 SBC 函数为：

$$\text{SBC}=n\ln(\hat{\sigma}_\varepsilon^2)+\ln(n)(p+q+1)$$

非中心化 $ARMA$ (p，q) 模型的 SBC 函数为：

$$\text{SBC}=n\ln(\hat{\sigma}_\varepsilon^2)+\ln(n)(p+q+2)$$

在所有通过检验的模型中使得 AIC 或 SBC 函数达到最小的模型为相对最优模型。之所以称为相对最优模型而不是绝对的最优模型，是因为我们不可能比较所有模型的 AIC 和 SBC 函数值，我们总是在尽可能全面的范围里考察有限多个模型的 AIC 和 SBC 函数值，再选择其中 AIC 和 SBC 函数值达到最小的那个模型作为最终的拟合模型，因而这样得到的最优模型就是一个相对最优模型。

【例 3.13 续】 用 AIC 准则和 SBC 准则评判例 3.13 中两个拟合模型的相对优劣。

检验结果如表 3—12 所示。

表 3—12

模型	AIC	SBC
$MA(2)$	536.455 6	543.201 1
$AR(1)$	535.789 6	540.286 6

最小信息量检验显示无论是使用 AIC 准则还是使用 SBC 准则，$AR(1)$模型都要优于 $MA(2)$模型，所以本例中 $AR(1)$模型是相对优化模型。

AIC 准则和 SBC 准则的提出，可以有效弥补根据自相关图和偏自相关图定阶的主观性，在有限的阶数范围内，帮助我们寻找相对最优拟合模型。

3.4 序列预测

到目前为止，我们对观察值序列做了许多工作，包括平稳性判别、白噪声判别、模型选择、参数估计及模型检验。这些工作的最终目的常常就是要利用这个拟合模型对随机序列的未来发展进行预测。

所谓预测就是要利用序列已观测到的样本值对序列在未来某个时刻的取值进行估计。目前对平稳序列最常用的预测方法是线性最小方差预测。线性是指预测值为观察值序列的线性函数，最小方差是指预测方差达到最小。

3.4.1 线性预测函数

根据 $ARMA(p,q)$ 模型的平稳性和可逆性，可以用传递形式和逆转形式等价描述该序列：

$$x_t = \sum_{i=0}^{\infty} G_i \varepsilon_{t-i} \tag{3.49}$$

$$\varepsilon_t = \sum_{j=0}^{\infty} I_j x_{t-j} \tag{3.50}$$

式中，$\{G_i\}$ 为 Green 函数值；$\{I_i\}$ 为逆转函数值。

把式（3.49）代入式（3.50），有

$$\begin{aligned} x_t &= \sum_{i=0}^{\infty} G_i \left(\sum_{j=0}^{\infty} I_j x_{t-i-j} \right) \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} G_i I_j x_{t-i-j} \end{aligned}$$

显然 x_t 是历史数据 x_{t-1}，x_{t-2}，…的线性函数。不妨简记为：

$$x_t = \sum_{i=0}^{\infty} C_i x_{t-1-i}$$

对任意一个未来时刻 $t+l$，$\forall l \geq 1$ 而言，该时刻的序列值 x_{t+l} 也可以表示成它的历史数据 x_{t+l-1}，…，x_{t+1}，x_t，x_{t-1}，…的线性函数：

$$x_{t+l} = \sum_{i=0}^{\infty} C_i x_{t+l-1-i}$$

但问题是其中只有部分历史信息 x_t，x_{t-1}，…是已知的，还有部分历史信息是未知的 x_{t+l-1}，…，x_{t+1}。

一个有趣的现象是，根据线性函数的可加性，所有的未知历史信息

x_{t+l-1}，…，x_{t+1}都可以用已知历史信息 x_t，x_{t-1}，…的线性函数表示出来。以 x_{t+2}为例，已知

$$x_{t+2}=\sum_{i=0}^{\infty}C_i x_{t+1-i}=C_0 x_{t+1}+\sum_{i=0}^{\infty}C_{i+1}x_{t-i} \tag{3.51}$$

式中，x_{t+1}是未知信息。

把 $x_{t+1}=\sum_{i=0}^{\infty}C_i x_{t-i}$ 代入式（3.51），得

$$\begin{aligned}x_{t+2}&=C_0\sum_{i=0}^{\infty}C_i x_{t-i}+\sum_{i=0}^{\infty}C_{i+1}x_{t-i}\\&=\sum_{i=0}^{\infty}(C_0C_i+C_{i+1})x_{t-i}\end{aligned}$$

由此 x_{t+2}最终表示成为已知历史信息 x_t，x_{t-1}，…的线性函数。

同理，对于未来任意 l 时刻的序列值 x_{t+l}，$\forall\, l\geqslant 1$ 最终都可以表示成已知历史信息 x_t，x_{t-1}，…的线性函数，并用该函数形式估计 x_{t+l}的值：

$$\hat{x}_t(l)=\sum_{i=0}^{\infty}\hat{D}_i x_{t-i} \tag{3.52}$$

$\hat{x}_t(l)$也称为序列 $\{x_t\}$ 的第 l 步预测值。

3.4.2 预测方差最小原则

用 $e_t(l)$衡量预测误差：

$$e_t(l)=x_{t+l}-\hat{x}_t(l)$$

显然，预测误差越小，预测精度就越高。因此，目前最常用的预测原则是预测方差最小原则，即

$$\text{Var}_{\hat{x}_t(l)}[e_t(l)]=\min\{\text{Var}[e_t(l)]\}$$

因为 $\hat{x}_t(l)$为 x_t，x_{t-1}，…的线性函数，所以该原则也称为线性预测方差最小原则。

为了便于分析，使用传递形式来描述序列值，根据 $ARMA(p,q)$平稳模型的性质和线性函数的可加性，显然有

$$\begin{cases}x_{t+l}=\sum_{i=0}^{\infty}G_i\varepsilon_{t+l-i}\\ \hat{x}_t(l)=\sum_{i=0}^{\infty}D_i x_{t-i}=\sum_{i=0}^{\infty}D_i\left(\sum_{j=0}^{\infty}G_j\varepsilon_{t-i-j}\right)=\sum_{i=0}^{\infty}W_i\varepsilon_{t-i}\end{cases}$$

则

$$e_t(l)=x_{t+l}-\hat{x}_t(l)$$

$$= \sum_{i=0}^{\infty} G_i \varepsilon_{t+l-i} - \sum_{i=0}^{\infty} W_i \varepsilon_{t-i} = \sum_{i=0}^{l-1} G_i \varepsilon_{t+l-i} - \sum_{i=0}^{\infty} (G_{l+i} - W_i) \varepsilon_{t-i}$$

预测方差为：

$$\mathrm{Var}[e_t(l)] = \left[\sum_{i=0}^{l-1} G_i^2 + \sum_{i=0}^{\infty} (G_{l+i} - W_i)^2\right]\sigma_\varepsilon^2 \geqslant \sum_{i=0}^{l-1} G_i^2 \sigma_\varepsilon^2$$

显然，要使得预测方差达到最小，必须要

$$W_i = G_{l+i}, i = 0, 1, 2, \cdots$$

这时，x_{t+l}的预测值为：

$$\hat{x}_t(l) = \sum_{i=0}^{\infty} G_{l+i} \varepsilon_{t-i}, \forall l \geqslant 1$$

预测误差为：

$$e_t(l) = \sum_{i=0}^{l-1} G_i \varepsilon_{t+l-i}$$

由于 $\{\varepsilon_t\}$ 为白噪声序列，所以

$$E[e_t(l)] = 0$$

$$\mathrm{Var}[e_t(l)] = \sum_{i=0}^{l-1} G_i^2 \sigma_\varepsilon^2, \forall l \geqslant 1$$

3.4.3　线性最小方差预测的性质

一、条件无偏最小方差估计偏

序列值 x_{t+l}可以如下分解：

$$\begin{aligned} x_{t+l} &= (\varepsilon_{t+l} + G_1 \varepsilon_{t+l-1} + \cdots + G_{l-1} \varepsilon_{t+1}) + (G_l \varepsilon_t + G_{l+1} \varepsilon_{t-1} + \cdots) \\ &= \qquad\qquad e_t(l) \qquad\qquad + \qquad \hat{x}_t(l) \end{aligned}$$

由式（3.52），$\hat{x}_t(l)$等价于

$$\hat{x}_t(l) = \sum_{i=0}^{\infty} \hat{D}_i x_{t-i}$$

即在 x_t，x_{t-1}，$\cdots$已知的条件下，$\hat{x}_t(l)$为常数，有

$$E(\hat{x}(l) \mid x_t, x_{t-1}, \cdots) = \hat{x}(l), \mathrm{Var}(\hat{x}(l) \mid x_t, x_{t-1}, \cdots) = 0$$

推导出

$$E(x_{t+l} \mid x_t, x_{t-1}, \cdots) = E[e_t(l) \mid x_t, x_{t-1}, \cdots] + E[\hat{x}(l) \mid x_t, x_{t-1}, \cdots] = \hat{x}(l)$$

$$\begin{aligned} \mathrm{Var}(x_{t+l} \mid x_t, x_{t-1}, \cdots) &= \mathrm{Var}[e_t(l) \mid x_t, x_{t-1}, \cdots] + \mathrm{Var}[\hat{x}(l) \mid x_t, x_{t-1}, \cdots] \\ &= \mathrm{Var}[e_t(l)] \end{aligned}$$

这说明在预测方差最小原则下得到的估计值$\hat{x}_t(l)$是序列值x_{t+l}在x_t，x_{t-1}，$\cdots$已知的情况下得到的条件无偏最小方差估计值。且预测方差只与预测步长 l 有

关，而与预测起始点 t 无关。但预测步长 l 越大，预测值的方差也越大，因而为了保证预测的精度，时间序列数据通常只适合做短期预测。

在正态假定下，有

$$x_{t+l}|x_t,x_{t-1},\cdots\sim N(\hat{x}_t(l),\mathrm{Var}[e_t(l)])$$

式中，$x_{t+l}\mid x_t$，x_{t-1}，…的置信水平为 $1-\alpha$ 的置信区间为：

$$(\hat{x}_t(l)\mp z_{1-\frac{\alpha}{2}}\cdot(1+G_1^2+\cdots+G_{l-1}^2)^{\frac{1}{2}}\cdot\sigma_\varepsilon)$$

二、$AR(p)$序列预测

在 $AR(p)$序列场合：

$$\begin{aligned}\hat{x}(l)&=E(x_{t+l}|x_t,x_{t-1},\cdots)\\&=E(\phi_1x_{t+l-1}+\cdots+\phi_px_{t+l-p}+\varepsilon_{t+l}|x_t,x_{t-1},\cdots)\\&=\phi_1\hat{x}_t(l-1)+\cdots+\phi_p\hat{x}_t(l-p)\end{aligned}$$

式中：

$$\hat{x}_t(k)=\begin{cases}\hat{x}_t(k),k\geqslant1\\x_{t+k}\quad,k\leqslant0\end{cases}$$

预测方差为：

$$\mathrm{Var}[e_t(l)]=(1+G_1^2+\cdots+G_{l-1}^2)\sigma_\varepsilon^2$$

【例 3.14】 已知某超市月销售额近似服从 $AR(2)$模型（单位：万元/每月）：

$$x_t=10+0.6x_{t-1}+0.3x_{t-2}+\varepsilon_1,\varepsilon_t\sim N(0,36)$$

今年第一季度该超市月销售额分别为：101 万元、96 万元、97.2 万元。请确定该超市第二季度每月销售额的 95%的置信区间。

（1）预测值的计算。

4 月份：$\hat{x}_3(1)=10+0.6x_3+0.3x_2=97.12$

5 月份：$\hat{x}_3(2)=10+0.6\hat{x}_3(1)+0.3x_3=97.432$

6 月份：$\hat{x}_3(3)=10+0.6\hat{x}_3(2)+0.3\hat{x}_3(1)=97.595\ 2$

（2）预测方差的计算。

首先，根据 Green 函数的递推公式，算得

$$G_0=1$$

$$G_1=\phi_1G_0=0.6$$

$$G_2=\phi_1G_1+\phi_2G_0=0.36+0.3=0.66$$

则

$$\mathrm{Var}[e_3(1)]=G_0^2\sigma_\varepsilon^2=36$$

$$\mathrm{Var}[e_3(2)]=(G_0^2+G_1^2)\sigma_\varepsilon^2=48.96$$

$$\text{Var}[e_3(3)]=(G_0^2+G_1^2+G_2^2)\sigma_\varepsilon^2=64.6416$$

(3) l 步预测销售额的95%置信区间为：

$$(\hat{x}_3(l)-1.96\sqrt{\text{Var}[e_3(l)]},\hat{x}_3(l)+1.96\sqrt{\text{Var}[e_3(l)]})$$

计算结果如表3—13所示。

表3—13

预测时期	95%置信区间
4月份	(85.36，108.88)
5月份	(83.72，111.15)
6月份	(81.84，113.35)

【例2.5续】 预测1999—2003年北京市城乡居民定期储蓄比例及相应的置信水平为95%的置信区间。

利用前面得到的拟合模型容易计算出以下结果，见表3—14。

表3—14

年份	预测值	95%置信区间下限	95%置信区间上限
1999	80.617 1	72.735 0	88.499 1
2000	80.905 5	71.322 9	90.488 0
2001	81.104 9	70.808 1	91.401 6
2002	81.242 7	70.621 5	92.863 9
2003	81.338 0	70.565 2	92.110 9

该序列拟合与预测图如图3—19所示。

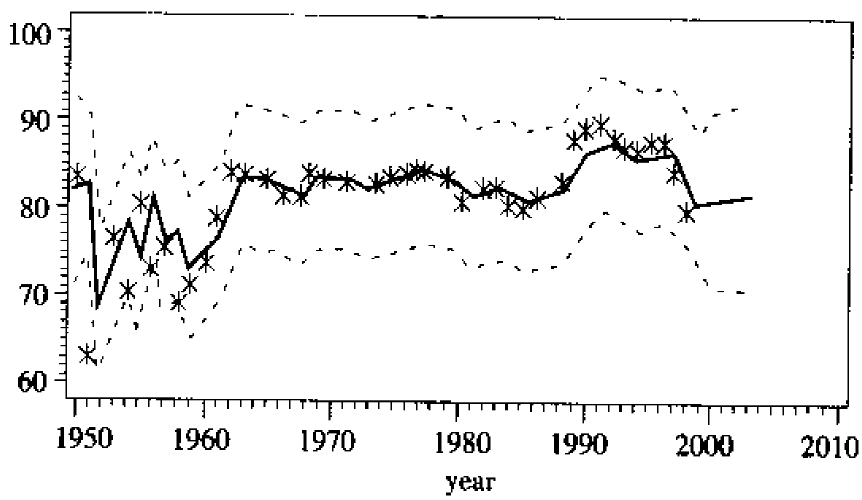


图3—19 北京市城乡居民定期储蓄比例序列拟合与预测图

图中，星号代表序列观察值；连续曲线代表拟合序列曲线；虚线代表拟合序列的95%上下置信限。可以看出随着预测期数的增加，预测方差也越来越大，

置信限呈现喇叭形。

三、MA(q)序列预测

对一个 $MA(q)$ 序列 $x_t=\mu+\varepsilon_t-\theta_1\varepsilon_{t-1}-\cdots-\theta_q\varepsilon_{t-q}$ 而言，有

$$x_{t+l}=\mu+\varepsilon_{t+l}-\theta_1\varepsilon_{t+l-1}-\cdots-\theta_q\varepsilon_{t+l-q}$$

在 x_t，x_{t-1}，…已知的条件下，求 x_{t+l} 的估计值，就等价于在 ε_t，ε_{t-1}，…已知的条件下，求 x_{t+l} 的估计值，而未来时刻的随机扰动 ε_{t+1}，ε_{t+2}，…是不可观测的，属于预测误差。所以：

当预测步长小于等于 MA 模型的阶数时($l\leqslant q$)，x_{t+l} 可以分解为：

$$\begin{aligned}x_{t+l}&=\mu+\varepsilon_{t+l}-\theta_1\varepsilon_{t+l-1}-\cdots-\theta_q\varepsilon_{t+l-q}\\&=(\varepsilon_{t+l}-\theta_1\varepsilon_{t+l-1}-\cdots-\theta_{l-1}\varepsilon_{t+1})+(\mu-\theta_l\varepsilon_t-\cdots-\theta_q\varepsilon_{t+l-q})\\&=e_t(l)+\hat{x}_t(l)\end{aligned}$$

当预测步长大于 MA 模型的阶数时($l>q$)，x_{t+l} 可以分解为：

$$\begin{aligned}x_{t+l}&=\mu+\varepsilon_{t+l}-\theta_1\varepsilon_{t+l-1}-\cdots-\theta_q\varepsilon_{t+l-q}\\&=(\varepsilon_{t+l}-\theta_1\varepsilon_{t+l-1}-\cdots-\theta_q\varepsilon_{t+l-q})+\mu\\&=e_t(l)+\hat{x}_t(l)\end{aligned}$$

即 $MA(q)$ 序列 l 步的预测值为：

$$\hat{x}_t(l)=\begin{cases}\mu-\sum\limits_{i=l}^{q}\theta_i\varepsilon_{t+l-i}, & l\leqslant q\\ \mu, & l>q\end{cases}$$

这说明 $MA(q)$ 序列理论上只能预测 q 步之内的序列走势，超过 q 步预报值恒等于序列均值。这是由 $MA(q)$ 序列自相关 q 步截尾的性质决定的。

$MA(q)$ 序列预测方差为：

$$\mathrm{Var}[e_t(l)]=\begin{cases}(1+\theta_1^2+\cdots+\theta_{l-1}^2)\sigma_\varepsilon^2, & l\leqslant q\\(1+\theta_1^2+\cdots+\theta_q^2)\sigma_\varepsilon^2, & l>q\end{cases}$$

【例 3.15】 已知某地区每年常住人口数量近似服从 $MA(3)$ 模型(单位:万人)：

$$x_t=10+\varepsilon_t-0.8\varepsilon_{t-1}+0.6\varepsilon_{t-2}-0.2\varepsilon_{t-1},\sigma_\varepsilon^2=25$$

最近 3 年的常住人口数量及一步预测数量如表 3—15 所示。

表 3—15

年份	统计人数	预测人数
2002	104	110
2003	108	100
2004	105	109

预测未来 5 年该地区常住人口的 95%置信区间。

（1）随机扰动项的计算：

$\varepsilon_{t-2}=x_{2002}-\hat{x}_{2001}(1)=104-110=-6$

$\varepsilon_{t-1}=x_{2003}-\hat{x}_{2002}(1)=108-100=8$

$\varepsilon_{t}=x_{2004}-\hat{x}_{2003}(1)=105-109=-4$

（2）未来常住人口预测值的计算：

$\hat{x}_t(1)=100-0.8\varepsilon_t+0.6\varepsilon_{t-1}-0.2\varepsilon_{t-2}=109.2$

$\hat{x}_t(2)=100+0.6\varepsilon_t-0.2\varepsilon_{t-1}=96$

$\hat{x}_t(3)=100-0.2\varepsilon_t=100.8$

$\hat{x}_t(4)=100$

$\hat{x}_t(5)=100$

（3）预测方差的计算：

$\mathrm{Var}[e_t(1)]=\sigma_\varepsilon^2=25$

$\mathrm{Var}[e_t(2)]=(1+\theta_1^2)\sigma_\varepsilon^2=41$

$\mathrm{Var}[e_t(3)]=(1+\theta_1^2+\theta_2^2)\sigma_\varepsilon^2=50$

$\mathrm{Var}[e_t(4)]=(1+\theta_1^2+\theta_2^2+\theta_3^2)\sigma_\varepsilon^2=51$

$\mathrm{Var}[e_t(5)]=(1+\theta_1^2+\theta_2^2+\theta_3^2)\sigma_\varepsilon^2=51$

（4）95%置信区间的计算：

$$\left(\hat{x}_t(l)-1.96\sqrt{\mathrm{Var}[e_t(l)]},\hat{x}_t(l)+1.96\sqrt{\mathrm{Var}[e_t(l)]}\right)$$

容易计算未来5年该地区常住人口95%的置信区间如表3－16所示。

表3—16

预测年份	95%置信区间
2005	(99, 119)
2006	(83, 109)
2007	(87, 115)
2008	(86, 114)
2009	(86, 114)

四、*ARMA*(*p*,*q*)序列预测

在 $ARMA(p,q)$ 模型场合：

$$\begin{aligned}x_t(l) &= E(\phi_1 x_{t+l-1}+\cdots+\phi_p x_{t+l-p}+\varepsilon_{t+l}-\theta_1\varepsilon_{t+l-1}\\&\quad-\cdots-\theta_q x_{t+l-q}\mid x_t,x_{t-1},\ldots)\\&=\begin{cases}\phi_1\hat{x}_t(l-1)+\cdots+\phi_p\hat{x}_t(l-p)-\sum\limits_{i=l}^{q}\theta_1\varepsilon_{t+l-i}, & l\leqslant q\\\phi_1\hat{x}_t(l-1)+\cdots+\phi_p\hat{x}_t(l-p), & l>q\end{cases}\end{aligned}$$

式中：

$$\tilde{x}_t(k)=\begin{cases}\hat{x}_t(k), k\geqslant 1\\ x_{t+k}\quad, k\leqslant 0\end{cases}$$

预测方差为：

$$\text{Var}[e_t(l)]=(G_0^1+G_1^2+\cdots+G_{l-1}^2)\sigma_\varepsilon^2$$

【例 3.16】 已知 *ARMA*（1，1）模型为：

$$x_t=0.8x_{t-1}+\varepsilon_t-0.6\varepsilon_{t-1}, \sigma_\varepsilon^2=0.002\,5$$

且 $x_{100}=0.3$，$\varepsilon_{100}=0.01$，预测未来 3 期序列值的 95%的置信区间。

（1）预测值的计算：

$$\hat{x}_{100}(1)=0.8x_{100}-0.6\varepsilon_{100}=0.234$$

$$\hat{x}_{100}(2)=0.8\hat{x}_{100}(1)=0.187\,2$$

$$\hat{x}_{100}(3)=0.8\hat{x}_{100}(2)=0.149\,76$$

（2）预测方差的计算：

首先，根据 Green 函数的递推公式，算得

$$G_0=1$$

$$G_1=\phi_1 G_0-\theta_1=0.2$$

$$G_2=\phi_1 G_1=0.16$$

则

$$\text{Var}[e_{100}(1)]=G_0^2\sigma_\varepsilon^2=0.002\,5$$

$$\text{Var}[e_{100}(2)]=(G_0^2+G_1^2)\sigma_\varepsilon^2=0.002\,6$$

$$\text{Var}[e_{100}(3)]=(G_0^2+G_1^2+G_2^2)\sigma_\varepsilon^2=0.002\,664$$

（3）95%置信区间的计算：

$$(\hat{x}_{100}(l)-1.96\sqrt{\text{Var}[e_{100}(l)]}, \hat{x}_{100}(l)+1.96\sqrt{\text{Var}[e_{100}(l)]})$$

计算结果见表 3—17。

表 3—17

时期	95%置信区间
101	(0.136，0.332)
102	(0.087，0.287)
103	(−0.049，0.251)

3.4.4 修正预测

对平稳时间序列的预测，实质就是根据所有的已知历史信息 x_t，x_{t-1}，…对序列未来某个时期的发展水平 x_{t+l}（$l=1$，2，…）作出估计。需要估计的时期

越长，我们的未知信息就越多。而未知信息越多，估计的精度就越差。

随着时间的发展，在原有观察值 x_t，x_{t-1}，…的基础上，我们会不断获得新的观察值 x_{t+1}，x_{t+2}，…。每获得一个新的观察值就意味着减少了一个未知信息，显然，如果能把新的信息加进来，就能够提高对 x_{t+l}的估计精度。所谓的修正预测就是研究如何利用新的信息去获得精度更高的预测值。

一个最简单的想法就是把新信息加入到旧的信息中，重新拟合模型，再利用拟合后的模型预测 x_{t+l}的序列值。在新的信息量比较大且使用统计软件很便利的时候，这不失为一种可行的修正办法。

但是在新的数据量不大或使用统计软件不是很方便的时候，这种重新拟合将是非常麻烦的一种修正方法。我们可以根据平稳时序预测的性质，寻找更为简便的修正预测方法。

已知在旧信息 x_t，x_{t-1}，…的基础上，x_{t-1}的预测值为：

$$\hat{x}_t(l)=G_l\varepsilon_t+G_1\varepsilon_{t-1}+\cdots$$

假如获得新的信息 x_{t+1}，则在 x_{t+1}，x_t，x_{t-1}，…的基础上，重新预测 x_{t+l}为：

$$\begin{aligned}\hat{x}_{t+1}(l-1)&=G_{l-1}\varepsilon_{t+1}+G_l\varepsilon_t+G_{l+1}\varepsilon_{t-1}+\cdots\\&=G_{l-1}\varepsilon_{t+1}+\hat{x}_t(l)\end{aligned}$$

式中，$\varepsilon_{t+1}=x_{t+1}-\hat{x}_t(l)$，是 x_{t+1}的一步预测误差。它的可测来源于 x_{t+1}提供的新信息。

此时，修正预测误差为：

$$e_{t+1}(l-1)=G_0\varepsilon_{t+l}+\cdots+G_{l-2}\varepsilon_{t+2}$$

因而，预测方差为：

$$\begin{aligned}\mathrm{Var}[e_{t+1}(l-1)]&=(G_0^2+\cdots+G_{l-2}^2)\sigma_\varepsilon^2\\&=\mathrm{Var}[e_t(l-1)]\end{aligned}$$

一期修正后第 l 步预测方差就等于修正前第 $l-1$ 步预测方差。它比修正前的同期预测方差减少了 $G_{l-1}^2\sigma_\varepsilon^2$，提高了预测精度。

上面的分析说明，当我们获得新的观察值时，要获得 x_{t+l}更精确的预测值并不需要重新对所有的历史进行计算。只需要利用新的观察值所带来的新的信息对旧的预测值进行修正即可。

更一般的情况，假如重新获得 p 个新观察值 x_{t+1}，…，x_{t+p}（$1\leqslant p\leqslant l$），则 x_{t+l}的修正预测值为：

$$\begin{aligned}\hat{x}_{t+p}(l-p)&=G_{l-p}\varepsilon_{t+p}+\cdots+G_{l-1}\varepsilon_{t+1}+G_l\varepsilon_t+G_{l+1}\varepsilon_{t-1}+\cdots\\&=G_{l-p}\varepsilon_{t+p}+\cdots+G_{l-1}\varepsilon_{t+1}+\hat{x}_t(l)\end{aligned}$$

式中，$\varepsilon_{t+i}=x_{t+i}-\hat{x}_{t+i-1}(1)$是 $x_{t+i}(i=1,2,\cdots,p)$的一步预测误差。

此时，修正预测误差为：

$$e_{t+p}(l-p)=G_0\varepsilon_{t+1}+\cdots+G_{l-p-1}\varepsilon_{t+p+1}$$

预测方差为：

$$\mathrm{Var}[e_{t+p}(l-p)]=(G_0^2+\cdots+G_{l-p-1}^2)\sigma_\varepsilon^2=\mathrm{Var}[e_t(l-p)]$$

【例 3.14 续】 假如一个月后知道 4 月份的真实销售额为 100 万元，求二季度后两个月销售额的修正预测值。

(1) 计算 4 月份销售额的一步预测误差：

$$\varepsilon_4=x_4-\hat{x}_3(1)=100-97.12=2.88$$

(2) 计算修正预测值。见表 3—18。

表 3—18

预测时期	预测值 $\hat{x}_3(l)$	新获得观察值	修正预测 $\hat{x}_4(l-1)$
1	97.12	100	
2	97.43		$\hat{x}_4(1)=G_1\varepsilon_4+\hat{x}_3(2)=99.16$
3	97.60		$\hat{x}_4(2)=G_2\varepsilon_4+\hat{x}_3(3)=99.50$

(3) 修正预测方差的计算：

$$\mathrm{Var}[e_4(1)]=\mathrm{Var}[e_3(1)]=G_0^2\sigma_\varepsilon^2=36$$

$$\mathrm{Var}[e_4(2)]=\mathrm{Var}[e_3(2)]=(G_0^2+G_1^2)\sigma_\varepsilon^2=48.96$$

(4) l 步预测销售额的 95%置信区间为：

$$(\hat{x}_4(l)-1.96\sqrt{\mathrm{Var}[e_4(l)]},\hat{x}_4(l)+1.96\sqrt{\mathrm{Var}[e_4(l)]})$$

计算结果见表 3—19。

表 3—19

预测时期	修正前置信区间	修正后置信区间
4 月份	(85.36，108.88)	
5 月份	(83.72，111.15)	(87.40，110.92)
6 月份	(81.84，113.35)	(85.79，113.21)

由修正前后的置信区间范围可以看出修正以后置信区间的宽度变小，即估计的精度提高了。

3.5 习 题

1. 已知 $AR(1)$ 模型为：$x_t=0.7x_{t-1}+\varepsilon_t$，$\varepsilon_t\sim WN(0,\sigma_\varepsilon^2)$。求 $E(x_t)$，

$\mathrm{Var}(x_t)$，ρ_2 和 ϕ_{22}。

2. 已知某 $AR(2)$模型为：$x_t=\phi_1 x_{t-1}+\phi_2 x_{t-2}+\varepsilon_t$，$\varepsilon_t \sim WN(0,\sigma_\varepsilon^2)$，且 $\rho_1=0.5$，$\rho_2=0.3$，求 ϕ_1，ϕ_2 的值。

3. 已知某 $AR(2)$模型为：$(1-0.5B)(1-0.3B)x_t=\varepsilon_t$，$\varepsilon_t \sim WN(0,\sigma_\varepsilon^2)$，求 $E(x_t)$，$\mathrm{Var}(x_t)$，ρ_k，ϕ_{kk}，其中 $k=1$，2，3。

4. 已知 $AR(2)$序列为 $x_t=x_{t-1}+cx_{t-1}+\varepsilon_t$，其中 $\{\varepsilon_t\}$ 为白噪声序列。确定 c 的取值范围，以保证 $\{x_t\}$ 为平稳序列，并给出该序列 ρ_k 的表达式。

5. 证明对任意常数 c，如下定义的 $AR(3)$序列一定是非平稳序列：

$$x_t=x_{t-1}+cx_{t-1}-xc_{t-2}+\varepsilon_t,\varepsilon_t \sim WN(0,\sigma_\varepsilon^2)$$

6. 对于 $AR(1)$模型：$x_t=\theta_1 x_{t-1}+\varepsilon_t,\varepsilon_t \sim WN(0,\sigma_\varepsilon^2)$，判断如下命题是否正确：

(1) $\gamma_0=(1+\theta_1^2)\sigma_\varepsilon^2$

(2) $E[(x_t-\mu)(x_{t-1}-\mu)]=-\theta_1$

(3) $\hat{x}_T(l)=\phi_1^l x_T$

(4) $e_T(l)=\varepsilon_{T+l}+\phi_1^2\varepsilon_{T+l-1}+\phi_1^4\varepsilon_{T+l-2}+\cdots+\phi_1^{2l}\varepsilon_T$

(5) $\lim\limits_{t\to\infty}\mathrm{Var}[x_{T+l}-\hat{x}_T(l)]=\frac{\sigma_\varepsilon^2}{\phi_1^2}$

7. 已知某中心化 $MA(1)$模型 1 阶自相关系数 $\rho_1=0.5$，求该模型的表达式。

8. 确定常数 C 的值，以保证如下表达式为 $MA(2)$模型：

$$x_t=10+0.5x_{t-1}+\varepsilon_t-0.8\varepsilon_{t-2}+C\varepsilon_{t-3}$$

9. 已知 $MA(2)$模型为：$x_t=\varepsilon_t-0.7\varepsilon_{t-1}+0.4\varepsilon_{t-2}$，$\varepsilon_t \sim WN(0,\sigma_\varepsilon^2)$。求 $E(x_t)$，$\mathrm{Var}(x_t)$，及 ρ_k $(k\geqslant 1)$。

10. 证明：

(1) 对任意常数 C，如下定义的无穷阶 MA 序列一定是非平稳序列：

$$x_t=\varepsilon_t+C(\varepsilon_{t-1}+\varepsilon_{t-2}+\cdots),\varepsilon_t \sim WN(0,\sigma_\varepsilon^2)$$

(2) $\{x_t\}$ 的 1 阶差分序列一定是平稳序列，并求 $\{y_t\}$ 自相关系数表达式：

$$y_t=x_t-x_{t-1}$$

11. 检验下列模型的平稳性与可逆性，其中 $\{\varepsilon_t\}$ 为白噪声序列：

(1) $x_t=0.5x_{t-1}+1.2x_{t-2}+\varepsilon_t$　　(2) $x_t=1.1x_{t-1}-0.3x_{t-2}+\varepsilon_t$

(3) $x_t=\varepsilon_t-0.9\varepsilon_{t-1}+0.3\varepsilon_{t-2}$　　(4) $x_t=\varepsilon_t+1.3\varepsilon_{t-1}-0.4\varepsilon_{t-2}$

(5) $x_t=0.7x_{t-1}+\varepsilon_t-0.6\varepsilon_{t-1}$　　(6) $x_t=-0.8x_{t-1}+0.5x_{t-2}+\varepsilon_t-1.1\varepsilon_{t-1}$

12. 已知 $ARMA$（1，1）模型为：$x_t=0.6x_{t-1}+\varepsilon_t-0.3\varepsilon_{t-1}$，确定该模型的 Green 函数，使该模型可以等价表示为无穷 MA 阶模型形式。

13. 某 $ARMA$（2，2）模型为：$\Phi(B)x_t=3+\Theta(B)\varepsilon_t$，求 $E(x_t)$。其中：

$\varepsilon_t \sim WN(0,\sigma_\varepsilon^2)$，$\Phi(B)=(1-0.5B)^2$。

14. 证明 $ARMA$（1，1）序列 $x_t=0.5x_{t-1}+\varepsilon_t-0.25\varepsilon_{t-1}$，$\varepsilon_t \sim WN(0,\sigma_\varepsilon^2)$ 的自相关系数为：

$$\rho_k=\begin{cases}1, & k=0\\ 0.27, & k=1\\ 0.5\rho_{k-1}, & k\geqslant 2\end{cases}$$

15. 对于平稳时间序列，如下等式哪些一定成立？

(1) $\sigma_\varepsilon^2=E(\varepsilon_1^2)$

(2) $\mathrm{Cov}(y_t,y_{t+k})=\mathrm{Cov}(y_t,y_{t-k})$

(3) $\rho_k=\rho_{-k}$

(4) $\hat{y}_t(k+1)=\hat{y}_{t+1}(k)$

16. 对于 $AR(1)$ 模型：$x_t-\mu=\phi_1(X_{t-1}-\mu)+\varepsilon_t$，根据 t 个历史观察值数据…,10.1，9.6 已求出 $\hat{\mu}=10$，$\hat{\phi}_1=0.3$，$\hat{\sigma}_\varepsilon^2=9$，求：

(1) x_{t+3} 的 95%的置信区间。

(2) 假定新获得观察值数据 $x_{t+1}=10.5$，用更新数据求 x_{t+3} 的 95%的置信区间。

17. 某城市过去 63 年中每年降雪量数据如表 3—20 所示。

表 3—20

126.4	82.4	78.1	51.1	90.9	76.2	104.5	87.4
110.5	25	69.3	53.5	39.8	63.6	46.7	72.9
79.6	83.6	80.7	60.3	79	74.4	49.6	54.7
71.8	49.1	103.9	51.6	82.4	83.6	77.8	79.3
89.6	85.5	58	120.7	110.5	65.4	39.9	40.1
88.7	71.4	83	55.9	89.9	84.8	105.2	113.7
124.7	114.5	115.6	102.4	101.4	89.8	71.5	70.9
98.3	55.5	66.1	78.4	120.5	97	110	

(1) 判断该序列的平稳性与纯随机性。

(2) 如果序列平稳且非白噪声，选择适当模型拟合该序列的发展。

(3) 利用拟合模型，预测该城市未来 5 年的降雪量。

18. 某地区连续 74 年的谷物产量（单位：千吨）如表 3—21 所示。

表 3—21

0.97	0.45	1.61	1.26	1.37	1.43	1.32	1.23	0.84	0.89	1.18
1.33	1.21	0.98	0.91	0.61	1.23	0.97	1.10	0.74	0.80	0.81
0.80	0.60	0.59	0.63	0.87	0.36	0.81	0.91	0.77	0.96	0.93
0.95	0.65	0.98	0.70	0.86	1.32	0.88	0.68	0.78	1.25	0.79

续前表

1.19	0.69	0.92	0.86	0.86	0.85	0.90	0.54	0.32	1.40	1.14
0.69	0.91	0.68	0.57	0.94	0.35	0.39	0.45	0.99	0.84	0.62
0.85	0.73	0.66	0.76	0.63	0.32	0.17	0.46			

（1）判断该序列的平稳性与纯随机性。

（2）选择适当模型拟合该序列的发展。

（3）利用拟合模型，预测该地区未来5年的谷物产量。

19. 现有201个连续的生产记录，如表3—22所示。

表3—22

81.9	89.4	79.0	81.4	84.8	85.9	88.0	80.3	82.6
83.5	80.2	85.2	87.2	83.5	84.3	82.9	84.7	82.9
81.5	83.4	87.7	81.8	79.6	85.8	77.9	89.7	85.4
86.3	80.7	83.8	90.5	84.5	82.4	86.7	83.0	81.8
89.3	79.3	82.7	88.0	79.6	87.8	83.6	79.5	83.3
88.4	86.6	84.6	79.7	86.0	84.2	83.0	84.8	83.6
81.8	85.9	88.2	83.5	87.2	83.7	87.3	83.0	90.5
80.7	83.1	86.5	90.0	77.5	84.7	84.6	87.2	80.5
86.1	82.6	85.4	84.7	82.8	81.9	83.6	86.8	84.0
84.2	82.8	83.0	82.0	84.7	84.4	88.9	82.4	83.0
85.0	82.2	81.6	86.2	85.4	82.1	81.4	85.0	85.8
84.2	83.5	86.5	85.0	80.4	85.7	86.7	86.7	82.3
86.4	82.5	82.0	79.5	86.7	80.5	91.7	81.6	83.9
85.6	84.8	78.4	89.9	85.0	86.2	83.0	85.4	84.4
84.5	86.2	85.6	83.2	85.7	83.5	80.1	82.2	88.6
82.0	85.0	85.2	85.3	84.3	82.3	89.7	84.8	83.1
80.6	87.4	86.8	83.5	86.2	84.1	82.3	84.8	86.6
83.5	78.1	88.8	81.9	83.3	80.0	87.2	83.3	86.6
79.5	84.1	82.2	90.8	86.5	79.7	81.0	87.2	81.6
84.4	84.4	82.2	88.9	80.9	85.1	87.1	84.0	76.5
82.7	85.1	83.3	90.4	81.0	80.3	79.8	89.0	83.7
80.9	87.3	81.1	85.6	86.6	80.0	86.6	83.3	83.1
82.3	86.7	80.2						

（1）判断该序列的平稳性与纯随机性。

（2）如果序列平稳且非白噪声，选择适当模型拟合该序列的发展。

（3）利用拟合模型，预测该序列下一时刻95%的置信区间。

3.6 上机指导

在 SAS/ETS 软件中有一个综合软件包——PROC ARIMA，在该程序中只需要输入非常简洁的命令就可以得到包括模型识别、参数估计、相对最优模型选择、短期预测等丰富的分析结果。

在第 2 章绘制序列自相关图时，我们简单介绍了 ARIMA 程序中的 IDENTIFY 命令，实际上一个完整的 ARIMA 程序是由 IDENTIFY（识别）、ESTIMATE（估计）和 FORECAST（预测）三条命令组成。这三条命令涵盖了平稳序列建模的每个步骤。它们既可以分开使用又可以联合使用。一个 ARIMA 程序可以包含多个 IDENTIFY，ESTIMATE 和 FORECAST 命令。

下面以临时数据集 example3_1 的数据为例详细介绍 ARIMA 程序的功能。首先建立该数据集，并绘制该序列时序图，时序图显示该序列波动平稳。

```
data example3_1;
input x@@;
time=_n_;
cards;
 0.30   -0.45   0.36   0.00   0.17   0.45    2.15
 4.42    3.48   2.99   1.74   2.40   0.11    0.96
 0.21   -0.10  -1.27  -1.45  -1.19  -1.47   -1.34
-1.02   -0.27   0.14  -0.07   0.10  -0.15   -0.36
-0.50   -1.93  -1.49  -2.35  -2.18  -0.39   -0.52
-2.24   -3.46  -3.97  -4.60  -3.09  -2.19   -1.21
 0.78    0.88   2.07   1.44   1.50   0.29   -0.36
-0.97   -0.30  -0.28   0.80   0.91   1.95    1.77
 1.80    0.56  -0.11   0.10  -0.56  -1.34   -2.47
 0.07   -0.69  -1.96   0.04   1.59   0.20    0.39
 1.06   -0.39  -0.16   2.07   1.35   1.46    1.50
 0.94   -0.08  -0.66  -0.21  -0.77  -0.52    0.05
;
proc gplot data=example3_1;
plot x*time=1;
symbol1 c=red I=join v=star;
run;
```

3.6.1 模型识别

一、IDENTIFY 语句介绍

ARIMA 过程的第一步是要使用 IDENTIFY 命令对该序列的平稳性和纯随机性进行识别，并对平稳非白噪声序列估计拟合模型的阶数。这些功能使用如下命令就可以实现。

```
proc arima data= example3_1;
identify Var=x nlag=8;
run;
```

每条 IDENTIFY 命令执行之后，会输出五方面的信息：

1. 分析变量的描述性统计

这部分输出的信息包括：

分析变量的名称；

该序列均值；

标准差及观察值的个数（样本容量）。

2. 样本自相关图

该图输出的信息上一章已经介绍过了。

3. 样本逆自相关图

该图从左到右输出的信息分别是：延迟阶数、逆自相关系数值和逆自相关图。

在前面我们详细介绍了样本自相关图和偏自相关图的概念与意义，还没有提到逆自相关的概念。所谓的逆自相关如下定义：

假定$\{x_t\}$服从平稳可逆 $ARMA(p,q)$过程

$$\Phi(B)x_t=\Theta(B)\varepsilon_t$$

式中，$\{\varepsilon_t\}$为白噪声序列。那么如下 $ARMA(q,p)$模型称为该 $ARMA(p,q)$模型的对偶模型：

$$\Theta(B)x_t=\Phi(B)\varepsilon_t$$

该对偶模型的自相关系数就称作原模型的逆自相关系数。

逆自相关系数和偏自相关系数有着基本相同的性质，但是它对于过差分有着非常敏感的判断，所以在后面讲到非平稳序列通过差分运算实现平稳时，需要考察逆自相关图的性质。

4. 样本偏自相关图

该图从左到右输出的信息分别是：延迟阶数、偏自相关系数值和偏自相关图。

5. 纯随机检验结果

该部分输出的信息上一章已经介绍过了。

根据上述五方面的信息，我们可以对观察值序列如下三方面的性质进行识别：

1. 平稳性

根据样本自相关图是否具有短期相关性，可以直观判别序列的平稳性。

2. 纯随机性

通过考察最后输出的自相关系数白噪声检验结果，可以对该序列的纯随机性进行识别。

3. 模型识别

根据样本自相关图、偏自相关图及逆自相关图的性质，可以对平稳非白噪声序列拟合模型的阶数进行直观识别。

二、相对最优定阶

为了尽量避免因个人经验不足导致的模型识别问题，SAS 系统还提供了相对最优模型识别。只要我们在 IDENTIFY 命令中增加一个可选命令 MINIC，就可以获得一定范围内的最优模型定阶。

比如本例中，我们将 IDENTIFY 命令改写成如下形式

```
identify Var=x nlag=8 minic p= (0: 5) q= (0: 5);
run;
```

其中 MINIC 选项是指定 SAS 系统输出所有自相关延迟阶数小于等于 5，移动平均延迟阶数也小于等于 5 的 $ARMA(p,q)$ 模型的 BIC 信息量，并指出其中 BIC 信息量达到最小的模型的阶数，这实际上就是模型优化的过程。

本例中，增加 MINIC 选项后，增加输出的信息如图 3—20 所示。

Minimum Information Criterion

Lags	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	0.756 693	0.566 331	0.345 231	0.070 485	−0.340 69	−0.303 54
AR 1	−0.279 6	−0.227 96	−0.189 01	−0.185 61	−0.302 9	−0.261 15
AR 2	−0.232 93	−0.180 92	−0.139 8	−0.134 54	−0.251 15	−0.209 6
AR 3	−0.188 05	−0.135 8	−0.092 01	−0.082 75	−0.199 09	−0.157 53
AR 4	−0.237 86	−0.187 99	−0.175 94	−0.123 37	−0.173 14	−0.140 08
AR 5	−0.237 19	−0.214 21	−0.212 02	−0.172 87	−0.134 42	−0.089 9

Error series model: AR(8)

Minimum Table Value: BIC (0, 4) =−0.340 69

图 3—20　IDENTIFY 命令输出的最小信息量结果

最后一条信息显示，在自相关延迟阶数小于等于 5，移动平均延迟阶数也小于等于 5 的所有 $ARMA(p,q)$模型中，BIC 信息量相对最小的是 $ARMA(0,4)$模型，即 $MA(4)$模型。

3.6.2 参数估计

确定了拟合模型的阶数之后，下一步就是要估计模型中未知参数的值，以确定模型的口径，并对拟合好的模型进行显著性诊断。使用如下命令可以轻松实现这些功能。

```
estimate q=4;
run;
```

每条 ESTIMATE 命令执行之后，会输出五方面的信息：

一、未知参数的估计值

输出的未知参数估计结果见图 3—21。

Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr> | t |	Lag
MU	−0.001 387 1	0.344 14	−0.00	0.996 8	0
MA1，1	−0.917 84	0.089 19	−10.29	<.000 1	1
MA1，2	−0.832 00	0.119 31	−6.97	<.000 1	2
MA1，3	−0.598 06	0.119 06	−5.02	<.000 1	3
MA1，4	−0.623 17	0.089 45	−6.97	<.000 1	4

图 3—21 ESTIMATE 命令输出的未知参数估计结果

SAS 系统支持三种参数估计方法：最小二乘估计方法、条件最小二乘估计方法及极大似然方法。如果我们不特别指定参数估计的方法，系统默认的估计方法是条件最小二乘估计方法。

如果需要特别指定参数估计方法，只需要在 ESTIMATE 命令中增加 MATHED 选项，可供选择的估计方法有：

METHOD=ML（极大似然估计方法）

METHOD=ULS（最小二乘估计方法）

METHOD=CLS（条件最小二乘估计方法）

参数估计部分输出的结果从左到右分别为：

1. 参数名称，其中 MU 为常数项，在本例中即为均值，*MA*1，1，…，*MA*1，4 分别为 θ_1，…，θ_4。

2. 各参数估计值。

3. 各参数估计量的标准差。

4. 各参数 t 检验统计量的值。

5. t 统计量的 P 值。

6. 各参数对应的延迟阶数。

本例参数估计输出结果显示均值 MU 不显著（t 检验统计量的 P 值为 0.996 8），其他参数均显著（t 检验统计量的 P 值均小于 0.000 1），所以选择 NOINT 选项，除去常数项，再次估计未知参数的结果，即可输入第二条 ESTIMATE 命令：

```
estimate q=4 noint;
run;
```

参数估计部分输出结果如图 3—22 所示。

Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr> | t |	Lag
MA1，1	−0.917 80	0.088 62	−10.36	<.000 1	1
MA1，2	−0.831 98	0.118 33	−7.03	<.000 1	2
MA1，3	−0.597 89	0.118 29	−5.05	<.000 1	3
MA1，4	−0.623 14	0.088 88	−7.01	<.000 1	4

图 3—22　ESTIMATE 命令消除常数项之后的输出结果

显然四个未知参数均显著。

二、拟合统计量的值

这部分输出五个统计量的值，由上到下分别是方差估计值、标准差估计值、AIC 信息量、SBC 信息量及残差个数。

三、系数相关阵

四、残差自相关检验结果

这部分的输出格式和序列自相关系数白噪声检验部分的输出结果一样。本例中由于延迟各阶的 LB 统计量的 P 值均显著大于 α（$\alpha=0.05$），所以该拟合模型显著成立。

五、拟合模型的具体形式

拟合模型的具体形式如图 3—23 所示。

Model for variable x

No mean term in this model.

Moving Average Factors

Factor 1：1+0.917 8 B** (1)+0.831 98 B** (2)+0.597 89 B** (3)+0.623 14 B** (4)

图 3—23　ESTIMATE 命令输出的拟合模型形式

该输出形式等价于

$$x_t=(1+0.917\ 8B+0.831\ 98B^2+0.559\ 789B^3+0.623\ 14B^4)\varepsilon_t$$

或记为

$$x_t=\varepsilon_t+0.917\ 8\varepsilon_{t-1}+0.831\ 98\varepsilon_{t-2}+0.559\ 789\varepsilon_{t-3}+0.623\ 14\varepsilon_{t-4}$$

本例中没有常数项也没有自相关因子，假定一个 *ARMA* 模型既含有常数项 μ，又含有自相关因子 $\Phi(B)$与移动平均因子 $\Theta(B)$，该模型应该表示为：

$$x_t=\mu+\frac{\Theta(B)}{\Phi(B)}\varepsilon_t$$

3.6.3　序列预测

模型拟合好之后，还可以利用该模型对序列进行短期预测。预测命令如下：

```
forecast lead=5 id=time out=results;
run;
```

其中，lead 是指定预测期数；id 是指定身份标识；out 是指定预测后的结果存入某个数据集。

该命令运行后会输出如下信息，如图 3—24 所示。

Forecasts for variable x

Obs	Forecast	Std Error	95% Confidence Limits	
85	0.618 5	0.873 9	1.094 3	2.331 4
86	0.272 5	1.186 2	2.052 5	2.597 4
87	0.392 3	1.391 3	2.334 6	3.119 3
88	0.469 6	1.486 2	2.443 3	3.382 5
89	0.000 0	1.582 8	3.102 3	3.102 3

图 3—24　FORECAST 命令输出的预测结果

该输出结果从左到右分别为序列值的序号、预测值、预测值的标准差、95%的置信下限、95%的置信上限。

利用存储在临时数据集 RESULTS 里的数据，我们还可以绘制漂亮的拟合、预测图，相关命令如下：

```
proc gplot data=results;
```

```
plot x * time=1 forecast * time=2 l95 * time=3 u95 * time=3/overlay;
symbol1 c=black i=none v=star;
symbol2 c=red i=join v=none;
symbol3 c=green i=join v=none l=32;
run;
```

输出图像如图 3—25 所示。

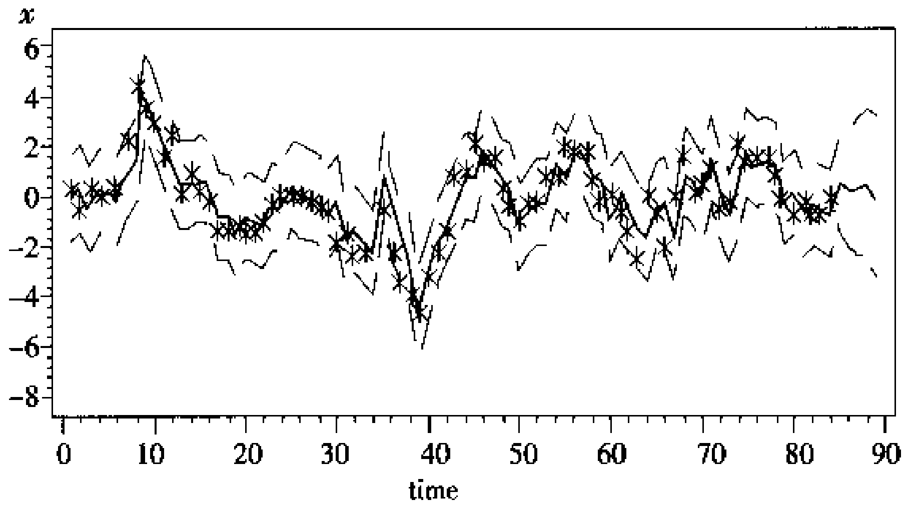


图 3—25　拟合效果图

第 4 章

非平稳序列的确定性分析

前面介绍了对平稳时间序列的分析方法。实际上，在自然界中绝大部分序列都是非平稳的，因而对非平稳序列的分析更普遍、更重要，人们创造的分析方法也更多。

对非平稳时间序列的分析方法可以分为确定性时序分析和随机时序分析两大类，本章主要介绍一些常用的确定性时序分析方法。

4.1 时间序列的分解

4.1.1 Wold 分解定理

1938 年，H. Wold 在他的博士论文“A Study in the Analysis of Stationary Time Series”中提出了著名的 Wold 分解定理。

Wold 分解定理 对于任何一个离散平稳过程 $\{x_t\}$，它都可以分解为两个不相关的平稳序列之和，其中一个为确定性的，另一个为随机性的，不妨记作

$$x_t = V_t + \xi_t$$

式中，$\{V_t\}$ 为确定性序列；$\{\xi_t\}$ 为随机序列，且 $\xi_t = \sum_{j=0}^{\infty}\varphi_j\varepsilon_{t-j}$。它们需要满足如下条件：

(1) $\varphi_0 = 1, \sum_{j=0}^{\infty}\varphi_j^2 < \infty$

(2) $\{\varepsilon_t\} \sim WN(0, \sigma_\varepsilon^2)$

(3) $E(V_t, \varepsilon_s) = 0, \forall t \neq S$

这里，所谓确定性序列和随机序列的概念如下定义：

对任意序列 $\{y_t\}$ 而言，令 y_t 关于 q 期之前的序列值 $\{y_{t-q}, y_{t-q-1}, \cdots\}$ 作线性回归

$$y_t = \alpha_0 + \alpha_1 y_{t-q} + \alpha_2 y_{t-q-1} + \cdots + \upsilon_t$$

式中，$\{\upsilon_t\}$ 为回归残差序列；$\mathrm{Var}(\upsilon_t) = \tau_q^2$。

显然，$\tau_q^2 \leqslant \mathrm{Var}(y_t)$，且随着 q 的增大而增大，也就是说 τ_q^2 是非减的有界序列，它的大小可以衡量历史信息对现时值的预测精度。τ_q^2 越小，说明预测得越准确；τ_q^2 越大，说明预测得越差。

如果 $\lim\limits_{q\to\infty}\tau_q^2 = 0$，即说明序列的发展有很强的规律性或者叫确定性，历史数据可以很好地预测将来，这时称 $\{y_t\}$ 是确定序列。

如果 $\lim\limits_{q\to\infty}\tau_q^2 = \mathrm{Var}(y_t)$，即说明序列随着时间的发展随机性很强，遥远的过去对于现时值的估计，或者说是预测效果很差，这时称 $\{y_t\}$ 是随机序列。

显然，之前我们对平稳序列进行 $ARMA(p, q)$ 模型拟合时，得到的拟合模型

$$x_t = \mu + \frac{\Theta(B)}{\Phi(B)}\varepsilon_t$$

其中，$\{\mu_t = \mu, t \in T\}$ 即为确定性平稳序列；$\left\{\frac{\Theta(B)}{\Phi(B)}\varepsilon_t, t \in T\right\}$ 即为随机平稳序列。而 $ARMA(p, q)$ 模型的分析实际上主要是对随机平稳序列进行的分析。

4.1.2 Cramer 分解定理

Wold 分解定理是现代时间序列分析理论的灵魂。尽管 Wold 提出这个分解定理只是为了分析平稳序列的构成，但 Cramer 已于 1961 年证明这种分解思路同样可以用于非平稳序列。

Cramer 分解定理 任何一个时间序列 $\{x_t\}$ 都可以分解为两部分的叠加：其中一部分是由多项式决定的确定性趋势成分，另一部分是平稳的零均值误差成

分，即

$$x_t = \mu_t + \varepsilon_t = \sum_{j=0}^{d} \beta_j t^j + \Psi(B) a_t$$

式中，$d<\infty$；β_1，…，β_d 为常数系数；$\{a_t\}$ 为一个零均值白噪声序列；B 为延迟算子。

因为

$$E(\varepsilon_t) = \Psi(B)E(a_t) = 0$$

所以有

$$E(x_t) = E(\mu_t) = \sum_{j=0}^{d} \beta_j t^j$$

即均值序列 $\left\{\sum_{j=0}^{d} \beta_j t^j\right\}$ 反映了 $\{x_t\}$ 受到的确定性影响，而$\{\varepsilon_t; \varepsilon_t = \Psi(B)a_t\}$反映了 $\{x_t\}$ 受到的随机影响。

Cramer 分解定理说明任何一个序列的波动都可以视为同时受到了确定性影响和随机性影响的综合作用。平稳序列要求这两方面的影响都是稳定的，而非平稳序列产生的机理就在于它所受到的这两方面的影响至少有一方面是不稳定的。

4.2 确定性因素分解

在自然界中，由确定性因素导致的非平稳，通常显示出非常明显的规律性，比如有显著的趋势或者有固定的变化周期，这种规律性信息通常比较容易提取，而由随机因素导致的波动则非常难以确定和分析。根据这种性质，传统的时序分析方法通常都把分析的重点放在确定性信息的提取上，忽视对随机信息的提取，通常将序列简单地假定为：

$$x_t = \mu_t + \varepsilon_t$$

式中，$\{\varepsilon_t\}$ 为零均值白噪声序列。这种分析方法就称为确定性分析方法。

最常用的确定性分析方法是确定性因素分解方法。该方法产生于长期的观察实践。人们发现尽管不同的序列的情况千变万化，但是序列的各种变化都可以归纳成四大类因素的综合影响：

1. 长期趋势。该因素的影响会导致序列呈现出明显的长期趋势（递增、递减等）。

2. 循环波动。该因素会导致序列呈现出从低到高再由高至低的反复循环波动。

3. 季节性变化。该因素会导致序列呈现出和季节变化相关的稳定的周期波动。

4. 随机波动。除了长期趋势、循环波动和季节性变化之外，序列还会受到各种其他因素的综合影响，而这些影响导致序列呈现出一定的随机波动。

在实际分析时，人们发现没有固定周期的循环波动与长期趋势的影响很难严格分解开，而有固定周期的循环波动和季节性变化又很难严格分解开。近年来，人们对四因素的确定性分析做了改进，现在通常把序列分解为三大因素的综合影响：

1. 长期趋势波动，它包括长期趋势和无固定周期的循环波动。

2. 季节性变化，它包括所有具有稳定周期的循环波动。

3. 随机波动，除了长期趋势波动和季节性变化之外，其他因素的综合影响归为随机波动。

这三大因素的综合影响会导致序列呈现出各种变化情况。而我们进行确定性时序分析的目的无外乎如下两种：

一是克服其他因素的影响，单纯测度出某一个确定性因素（长期趋势波动或季节效应）对序列的影响。

二是推断出各种确定性因素彼此之间的相互作用关系及它们对序列的综合影响。

4.3 趋势分析

有些时间序列具有非常显著的趋势，有时我们分析的目的就是要找到序列中的这种趋势，并利用这种趋势对序列的发展作出合理的预测。

4.3.1 趋势拟合法

趋势拟合法就是把时间作为自变量，相应的序列观察值作为因变量，建立序列值随时间变化的回归模型的方法。根据序列所表现出的线性或非线性特征，我们的拟合方法又可以具体分为线性拟合和曲线拟合。

一、线性拟合

如果长期趋势呈现出线性特征，那么我们可以用线性模型来拟合它，模型可以具体写为：

$$\begin{cases} x_t = a + bt + I_t \\ E(I_t) = 0, \mathrm{Var}(I_t) \end{cases}$$

式中，$\{I_t\}$ 为随机波动；$T_t = a + bt$ 就是消除随机波动的影响之后该序列的长期趋势。

【例 4.1】 澳大利亚政府 1981—1990 年每季度的消费支出数据（单位：百万美元）如表 4—1 所示。

表 4—1

8 444	9 215	8 879	8 990	8 115	9 457	8 590	9 294	8 997	9 574
9 051	9 724	9 120	10 143	9 746	10 074	9 578	10 817	10 116	10 779
9 901	11 266	10 686	10 961	10 121	11 333	10 677	11 325	10 698	11 624
11 052	11 393	10 609	12 077	11 376	11 777	11 225	12 231	11 884	12 109

该序列时序图显示该序列有显著的线性递增趋势。于是考虑使用线性模型

$$\begin{cases} x_t = a + bt + I_t, t = 1, 2, \cdots, 40 \\ E(I_t) = 0, \mathrm{Var}(I_t) = \sigma^2 \end{cases}$$

拟合该序列的发展。

使用最小二乘法得到未知参数的估计值为：

$\hat{a} = 8\ 498.69$，$\hat{b} = 89.12$

对拟合模型进行检验，检验结果显示方程显著成立，且参数显著非零。拟合效果图如图 4—1 所示。

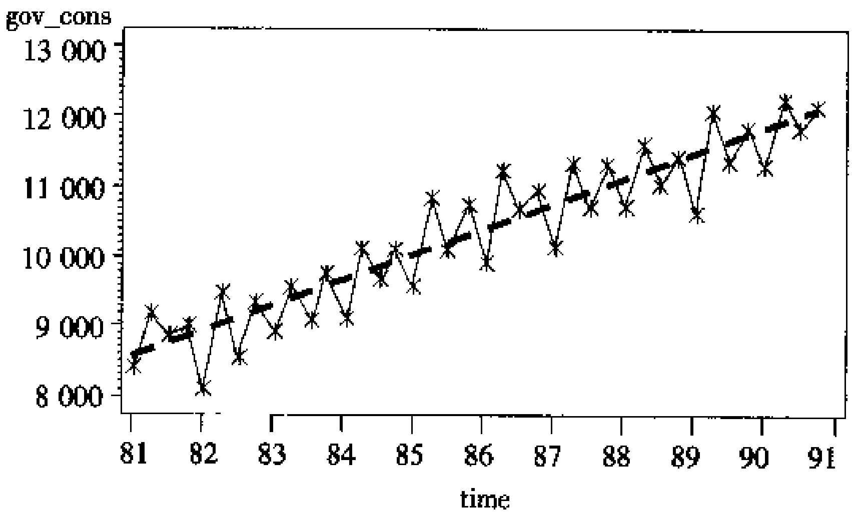


图 4—1 澳大利亚政府季度消费支出序列线性拟合图

图中，星号表示序列观察值；虚直线为线性趋势拟合线。

二、曲线拟合

如果长期趋势呈现出非线性特征，那么我们可以用曲线模型来拟合它。

对曲线模型进行参数估计时，指导思想是：能转换成线性模型的都转换成线性模型，用线性最小二乘法进行参数估计；实在不能转换成线性的，就用迭代法进行参数估计。

常用的曲线模型和对应的参数估计方法见表 4—2。

表 4—2

模型	变换	参数估计方法
二次型：$T_t=a+bt+ct^2$	令 $t_2=t^2$，原模型变换为：$T_t=a+bt+ct_2$	线性最小二乘法
指数型：$T_t=ab^t$	对原模型求对数，再令 $T_t'=\ln T_t$，$a'=\ln a$，$b'=\ln b$，原模型变换为：$T_t'=a'+b't$	用线性最小二乘法求出 a'，b'，再做变换：$a=e^{a'}$，$b=e^{b'}$
修正指数型：$T_t=a+bc^t$	不能转换成线性模型	迭代法
Gompertz 型：$T_t=e^{a+bc^t}$	不能转换成线性模型	迭代法
Logistic 型：$T_t=\dfrac{1}{a+bc^t}$	不能转换成线性模型	迭代法

【例 4.2】 对上海证券交易所 1991 年 1 月—2001 年 10 月每月末上证指数序列进行模型拟合（数据见附录 1.9）。

时序图显示该序列有显著的曲线递增趋势。尝试使用二次型模型

$$T_t=a+bt+ct^2,\quad t=1,2,\cdots,130$$

拟合该序列的发展。

1. 先做变换：把 t^2 的值赋给 t_2，原模型变换为线性模型：

$$T_t=a+bt+ct_2$$

2. 利用线性最小二乘法得到线性模型中未知参数的估计值：

$$\hat{a}=457.535\ 3,\ \hat{b}=1.811\ 9,\ \hat{c}=0.082\ 2$$

3. 检验方程。发现方程显著（p 值小于 0.000 1），但是参数 b 不显著（p 值为 0.450 5）。

4. 除去不显著自变量 t，拟合新的线性模型：

$$T_t=a+ct_2。$$

5. 求得该模型的未知参数的最小二乘估计值为：

$$\hat{a}=502.251\ 7,\ \hat{c}=0.095\ 2$$

6. 检验方程，方程及各参数均显著。

所以可以用二次型 $T_t=502.2517+0.0952t^2$ 拟合近 11 年来上证指数的长期变化趋势。拟合效果如图 4—2 所示。

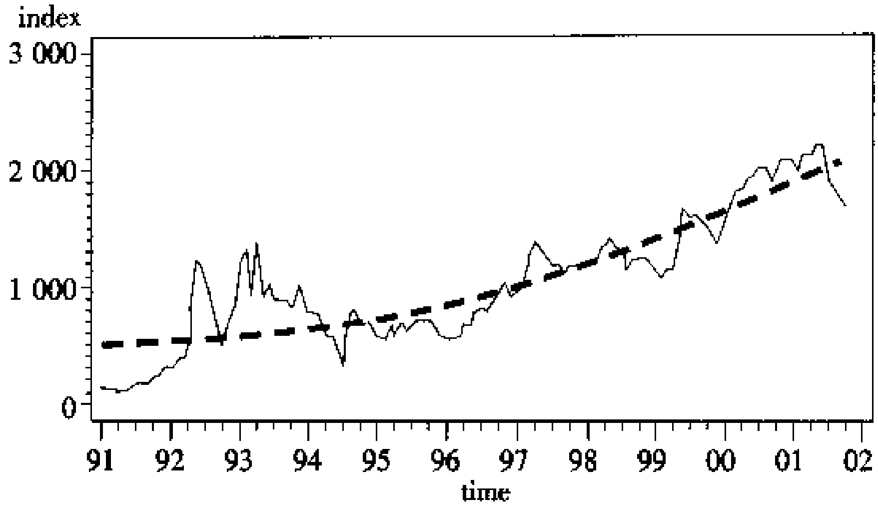


图 4—2　上证指数序列曲线拟合图

图中，连续曲线为观察值序列时序图；虚线为二次型拟合曲线。

4.3.2　平滑法

平滑法是进行趋势分析和预测时常用的一种方法。它是利用修匀技术，削弱短期随机波动对序列的影响，使序列平滑化，从而显示出变化的规律。它具有调节灵活、计算简便的特征，广泛应用于计量经济、人口研究等诸多领域。根据所用的平滑技术的不同，平滑法又可以具体分为移动平均法和指数平滑法。

一、移动平均法

移动平均法的基本思想是对于一个时间序列 $\{X_t\}$，我们可以假定在一个比较短的时间间隔里，序列的取值是比较稳定的，它们之间的差异主要是由随机波动造成的。根据这种假定，我们可以用一定时间间隔内的平均值作为某一期的估计值。它又可以分为两类：

1．n 期中心移动平均

$$\tilde{x}_t=\begin{cases}\frac{1}{n}\left(x_{t-\frac{n-1}{2}}+x_{t-\frac{n-1}{2}+1}+\cdots+x_t+\cdots+x_{t+\frac{n-1}{2}-1}+x_{t+\frac{n-1}{2}}\right), & n\text{ 为奇数}\\ \frac{1}{n}\left(\frac{1}{2}x_{t-\frac{n}{2}}+x_{t-\frac{n}{2}+1}+\cdots+x_t+\cdots+x_{t+\frac{n}{2}-1}+\frac{1}{2}x_{t+\frac{n}{2}}\right), & n\text{ 为偶数}\end{cases}$$

2．n 期移动平均

$$\tilde{x}_t=\frac{1}{n}(x_t+x_{t-1}+\cdots+x_{t-n+1})$$

移动平均的期数对原序列的修匀效果影响很大，要确定移动平均的期数，一般会从如下三个方面加以考虑：

（1）事件的发展有无周期性。如果事件的发展具有一定的周期性，一般以周期长度作为移动平均的间隔长度。比如研究每月的平均气温变化趋势，就应该做12期移动平均。通过周期平滑消除季节效应的影响。

（2）我们对趋势平滑性的要求。一般移动平均的期数越多，修匀曲线越平滑，表现出的长期趋势就越清晰。

（3）我们对趋势反映近期变化敏感程度的要求。用移动平均方法确定事件的发展趋势都具有一定的滞后性。移动平均的期数越多，滞后性越大，移动平均的期数越少，所得的趋势图对近期变化的反应就越敏感。

综合如上三方面的考虑，如果想得到长期趋势，就应该做期数较大的移动平均，如果想密切关注序列的短期趋势，就应该做期数较小的移动平均。

在预测领域，n 期移动平均还是一种常用的预测方法。假定最后一期的观察值为 x_T，那么使用 n 期移动平均方法，向前 l 期的预测值为：

$$\hat{x}_{T+l}=\frac{1}{n}(x'_{T+l-1}+x'_{T+l-2}+\cdots+x'_{T+l-n}),l\geqslant 1$$

式中：

$$x'_{T+l-i}=\begin{cases}\hat{x}_{T+l-i},l>i\\ x_{T+l-i},l\leqslant i\end{cases},i=1,2,\cdots,n$$

【例 4.3】 某一观察值序列 $\{x_t\}$ 最后 4 期的观察值 x_{T-3}，x_{T-2}，x_{T-1}，x_T 为：

5，5.4，5.8，6.2

（1）使用 4 期移动平均法预测 $\hat{x}_{T+2}$。

（2）求在 2 期预测值 $\hat{x}_{T+2}$ 中 x_T 前面的系数等于多少？

解：（1）$\hat{x}_{T+1}=\frac{1}{4}(x_T+x_{T-1}+x_{T-2}+x_{T-3})=\frac{5+5.4+5.8+6.2}{4}=5.6$

$$\hat{x}_{T+2}=\frac{1}{4}(\hat{x}_{T+1}+x_T+x_{T-1}+x_{T-2})=\frac{5.6+5+5.4+5.8}{4}=5.45$$

（2）因为

$$\begin{aligned}\hat{x}_{T+2}&=\frac{1}{4}(\hat{x}_{T+1}+x_T+x_{T-1}+x_{T-2})\\&=\frac{1}{4}\left[\frac{1}{4}(x_T+x_{T-1}+x_{T-2}+x_{T-2})+x_T+x_{T-1}+x_{T-2}\right]\\&=\frac{5}{16}(x_T+x_{T-1}+x_{T-2})+\frac{1}{16}x_{T-3}\end{aligned}$$

所以使用 4 期移动平均法，在 2 期预测值 $\hat{x}_{T+2}$ 中 x_T 前面的系数等于$\frac{5}{16}$。

二、指数平滑法

移动平均法实际上就是用一个简单的加权平均数作为某一期趋势的估计值。以 n 期移动平均为例，$\tilde{x}_t=\frac{1}{n}(x_t+x_{t-1}+\cdots+x_{t-n+1})$，相当于用近 n 期的加权平均数作为最后一期趋势的估计值，它们的权重都取为$\frac{1}{n}$，实际上也就是假定无论时间的远近，这 n 期的观察值 x_t，x_{t-1}，…，x_{t-n+1} 对最后一期的影响都是一样的。但在实际生活中，我们会发现对大多数随机事件而言，一般都是近期的结果对现在的影响会大些，远期的结果对现在的影响会小些。为了更好地反映这种影响作用，我们将考虑到时间间隔对事件发展的影响，各期权重随时间间隔的增大而呈指数衰减。这就是指数平滑法的基本思想。不考虑季节（周期）因素的影响，常用的进行趋势拟合的指数平滑公式有如下两种：

1. 简单指数平滑

$$\tilde{x}_t=\alpha x_t+\alpha(1-\alpha)x_{t-1}+\alpha(1-\alpha)^2x_{t-2}+\cdots$$

式中，α 为平滑系数，它满足 $0<\alpha<1$。

因为

$$\tilde{x}_{t-1}=\alpha x_{t-1}+\alpha(1-\alpha)x_{t-2}+\alpha(1-\alpha)^2x_{t-3}+\cdots$$

所以 $\tilde{x}_t$ 又等价于

$$\tilde{x}_t=\alpha x_t+(1-\alpha)\tilde{x}_{t-1}$$

简单指数平滑面临一个确定 $\tilde{x}_0$ 初始值的问题。我们有许多方法可以确定 $\tilde{x}_0$ 的初始值，最简单的方法是指定 $\tilde{x}_0=x_1$。

平滑系数 α 的值由研究人员根据经验给出。一般对于变化缓慢的序列，常取较小的 α 值，相反对于变化迅速的序列，常取较大的 α 值。经验表明 α 的值介于 0.05～0.3 之间，修匀效果比较好。

指数平滑法也是一种常用的预测方法，$\tilde{x}_T$ 常常作为 1 期预测值：

$$\begin{aligned}\hat{x}_{T+1}&=\tilde{x}_T\\&=\alpha x_T+\alpha(1-\alpha)x_{T-1}+\alpha(1-\alpha)^2x_{T-2}+\cdots\\&=\alpha x_T+(1-\alpha)\hat{x}_T\end{aligned}$$

指数平滑 2 期预测值为：

$$\begin{aligned}\hat{x}_{T+2}&=\alpha\hat{x}_{T+1}+\alpha(1-\alpha)x_T+\alpha(1-\alpha)^2x_{T-1}+\cdots\\&=\alpha\hat{x}_{T+1}+(1-\alpha)\hat{x}_{T+1}\\&=\hat{x}_{T+1}\end{aligned}$$

容易验证，指数平滑 l 期预测值都具有如下关系：

$$\hat{x}_{T+l}=\hat{x}_{T+1}，\ l\geqslant 2$$

【例 4.4】 对某一观察值序列 $\{x_t\}$ 使用指数平滑法。已知 $x_T=10$，$\tilde{x}_{T-1}=10.5$，平滑系数 $\alpha=0.25$。

(1) 求 2 期预测值 $\hat{x}_{T+2}$。

(2) 求在 2 期预测值 $\hat{x}_{T+2}$ 中 x_T 前面的系数等于多少？

解：(1) $\hat{x}_{T+1}=\tilde{x}_T=0.25x_T+0.75\tilde{x}_{T-1}=10.3$

$$\hat{x}_{T+2}=\hat{x}_{T-1}=10.3$$

(2) 因为

$$\hat{x}_{T+2}=\hat{x}_{T-1}=\alpha x_T+\alpha(1-\alpha)x_{T-1}+\cdots$$

所以使用指数平滑法在 2 期预测使 $\hat{x}_{T+2}$ 中 x_T 前面的系数就等于平滑系数 α，本例中 $\alpha=0.25$。

2. Holt 两参数指数平滑

Holt 两参数指数平滑适用于对含有线性趋势的序列进行修匀。它的基本思想是假定序列有一个比较固定的线性趋势——每期都递增 r 或递减 r，那么第 t 期的估计值就应该等于第 $t-1$ 期的观察值加上每期固定的趋势变动值，即

$$\hat{x}_t=x_{t-1}+r$$

但是由于随机因素的影响，使得每期的递增或递减值不会恒定为 r，它会随时间变化上下波动，所以趋势序列实际上是一个随机序列 $\{r_t\}$，因而

$$\hat{x}_t=x_{t-1}+r_{t-1}$$

考虑用第 t 期的观察值和第 t 期的估计值的加权平均数作为第 t 期的修匀值：

$$\begin{aligned}&\tilde{x}_t=\alpha x_t+(1-\alpha)\hat{x}_t\\&\Rightarrow\hat{x}_t=\alpha x_t+(1-\alpha)(x_{t-1}+r_{t-1})\end{aligned},0<\alpha<1 \tag{4.1}$$

因为趋势序列 $\{r_t\}$ 也是一个随机序列，为了让修匀序列 $\{\tilde{x}_t\}$ 更平滑，我们对 $\{r_t\}$ 也进行一次修匀处理：

$$r_t=\gamma(\tilde{x}_t-\tilde{x}_{t-1})+(1-\gamma)r_{t-1},0<\gamma<1 \tag{4.2}$$

把式 (4.2) 代入式 (4.1)，就能得到比较光滑的修匀序列 $\{\tilde{x}_t\}$。这就是 Holt 两参数指数平滑法的构造思想，它的平滑公式为：

$$\begin{cases}\tilde{x}_t=\alpha x_t+(1-\alpha)(\tilde{x}_{t-1}+r_{t-1})\\ r_1=\gamma(\tilde{x}_t-\tilde{x}_{t-1})+(1-\gamma)r_{t-1}\end{cases}$$

式中，α，γ 为两个平滑系数，也称为两个平滑参数，它们满足 $0<\alpha$，$\gamma<1$。平滑系数的选择原则和简单指数平滑的原则一样。

和简单指数平滑一样，我们也面临确定初始值的问题，在此我们需要确定两

个序列的初始值：

（1）平滑序列的初始值 $\tilde{x}_0$。最简单的是指定 $\tilde{x}_0=x_1$。

（2）趋势序列的初始值 r_0。和确定 $\tilde{x}_0$ 一样，我们有许多方法确定它，最简单的方法是：任意指定一个区间长度 n，用这段区间的平均趋势作为趋势初始值：

$$r_0=\frac{x_{n+1}-x_1}{n}$$

假定最后一期的修匀值为 $\tilde{x}_T$，那么使用 Holt 两参数指数平滑方法，向前 l 期的预测值为：

$$\hat{x}_{T+1}=\tilde{x}_T+l\cdot r_T$$

【例 4.5】 对北京市 1978—2000 年报纸发行量序列进行 Holt 两参数指数平滑。指定 $\tilde{x}_0=x_1=51\ 259$，$r_0=\frac{x_{23}-x_1}{23}=4\ 325$，$\alpha=0.15$，$\gamma=0.1$。

根据各年度报纸发行量计算出的 Holt 两参数指数平滑估计值（单位：万份/年）如表 4—3 所示。

表 4—3

年份	报纸发行量（万份/年）	Holt 两参数指数平滑值（$\alpha=0.15$，$\gamma=0.1$）
1978	51 259	43 570.52
1979	63 565	44 070.6
1980	75 095	46 449.4
1981	78 371	50 457.61
1982	78 984	54 725.55
1983	86 499	58 799.93
1984	98 628	63 693.16
1985	99 941	70 013.85
1986	103 630	76 012.59
1987	109 547	82 023.81
1988	113 375	88 344.98
1989	82 999	94 610.2
1990	87 489	95 660.72
1991	94 339	97 037.23
1992	114 824	99 091.54
1993	127 791	103 839.2
1994	114 373	109 984.5
1995	112 577	113 462.4

续前表

年份	报纸发行量 （万份/年）	Holt 两参数指数平滑值 （α=0.15，γ=0.1）
1996	136 308	116 162.9
1997	130 754	121 964.1
1998	140 870	126 277.3
1999	148 039	131 528
2000	146 395	137 206.6

平滑曲线如图 4—3 所示。

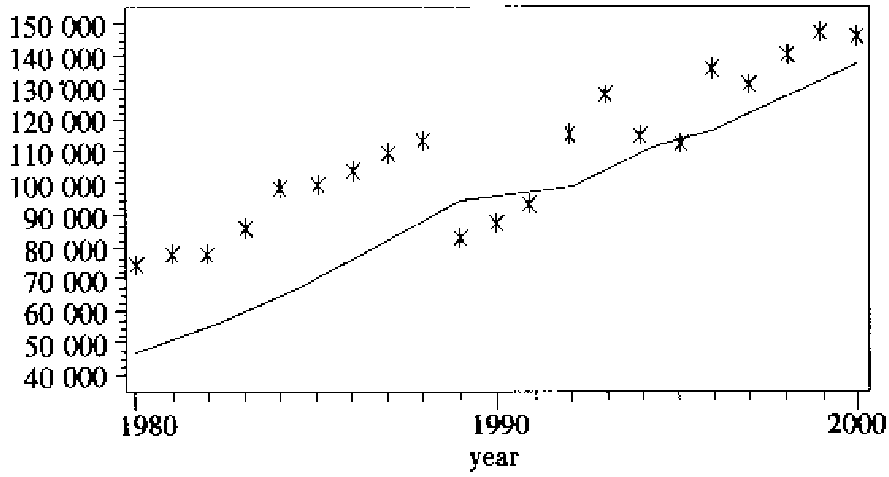


图 4—3　Holt 两参数指数平滑时序图

图中，星号为原序列观察值；曲线为 Holt 两参数指数平滑曲线。

4.4　季节效应分析

在日常生活中，我们可以见到许多有季节效应的时间序列，比如：四季的气温、每个月的商品零售额、某自然景点每季度的旅游人数等等。它们都会呈现出明显的季节变动规律。

我们还可以把“季节”广义化，凡是呈现出固定的周期性变化的事件，我们都称它具有“季节”效应。现在“季节”效应已经变成周期效应的代名词。而“季”也变成周期内每一期的代名词。

【例 4.6】　以北京市 1995—2000 年月平均气温序列为例（数据见附录 1.10），介绍季节效应分析的基本思想和具体操作步骤。

首先绘制该序列的时序图，如图 4—4 所示。

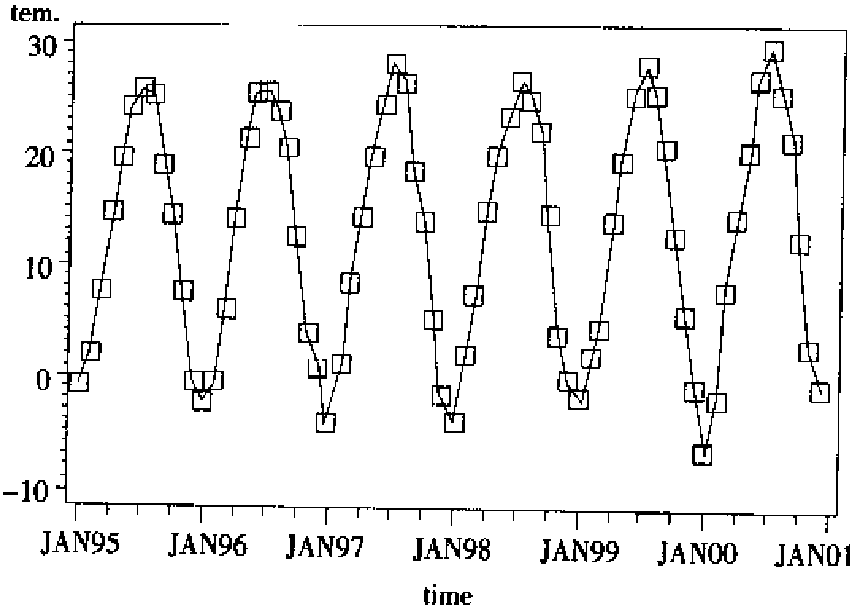


图 4—4　北京市 1995—2000 年月平均气温序列时序图

通过时序图，我们发现北京市 1995—2000 年每月的平均气温随着季节的变动有着非常规律的变化。气温的波动主要受到两个因素的影响：一个是季节效应；一个是随机波动。

假如没有季节效应的影响，北京市的气温应该始终在某个均值附近随机波动，因为季节效应的存在，使得气温会在不同年份的相同月份呈现出相似的性质，为了便于得到数量化的季节信息，我们构造出季节指数的概念。所谓季节指数就是用简单平均法计算的周期内各时期季节性影响的相对数。

以本例的数据结构为例，我们需要求出每个月的季节指数 S_1，S_2，…，S_{12}，那么第 i 年第 j 个月的平均气温可以表示为：

$$x_{ij}=\bar{x}\cdot S_j+I_{ij},\ j=1,\ 2,\ \cdots,\ 12$$

式中，$\bar{x}$ 为各月总平均气温；S_j 为第 j 个月的季节指数。I_{ij} 为第 i 年第 j 个月气温的随机波动。

季节指数的计算分为三步：

第一步：计算周期内各期平均数，得到长期以来该时期的平均水平。假定序列的数据结构为 m 期为一周期，共有 n 个周期。则

$$\bar{x}_k=\frac{\sum_{i=1}^{n}x_{ik}}{n},k=1,2,\cdots,m$$

第二步：计算总平均数。

$$\bar{x}=\frac{\sum_{i=1}^{n}\sum_{k=1}^{m}x_{ik}}{nm}$$

第三步：用时期平均数除以总平均数就可以得到各时期的季节指数 S_k（$k=1, 2, \cdots, m$），即

$$S_k=\frac{\bar{x}_k}{\bar{x}},\ k=1,\ 2,\ \cdots,\ m$$

这就是季节指数的构造方法。季节指数反映了该季度与总平均值之间的一种比较稳定的关系。如果这个比值大于 1，就说明该季度的值常常会高于总平均值；如果这个比值小于 1，就说明该季度的值常常低于总平均值；如果序列的季节指数都近似等于 1，那就说明该序列没有明显的季节效应。

下面我们来具体计算本例的季节指数。计算结果如表 4—4 所示。

表 4—4

年 月	平均气温（℃）						月平均 $\bar{x}_j$	季节指数 S_j
	1995	1996	1997	1998	1999	2000		
1	0.7	−2.2	−3.8	−3.9	−1.6	−6.4	−3.10	−0.238
2	2.1	−0.4	1.3	2.4	2.2	−1.5	1.02	0.078
3	7.7	6.2	8.7	7.6	4.8	8.1	7.18	0.551
4	14.7	14.3	14.5	15.0	14.4	14.6	14.58	1.119
5	19.8	21.6	20.0	19.9	19.5	20.4	20.20	1.550
6	24.3	25.4	24.6	23.6	25.4	26.7	25.00	1.919
7	25.9	25.5	28.2	26.5	28.1	29.6	27.30	2.095
8	25.4	23.9	26.6	25.1	25.6	25.7	25.38	1.948
9	19.0	20.7	18.6	22.2	20.9	21.8	20.53	1.576
10	14.5	12.8	14.0	14.8	13.0	12.6	13.62	1.045
11	7.7	4.2	5.4	4.0	5.9	3.0	5.03	0.386
12	−0.4	0.9	−1.5	0.1	−0.6	−0.6	−0.35	−0.027
总平均 $\bar{x}=13.03$								

把季节指数绘制成图，可以帮助我们更清楚地看到月度变迁对气温的影响。

本例中，7 月份的季节指数最大，说明 7 月份是北京最热的月份；1 月份的季节指数最小，说明 1 月份是北京最冷的月份；最接近于年平均气温的是 10 月份，这个月的平均气温是 13.62℃，和年平均气温的差异只有 4.5%。

由任意两个季节指数，我们可以得到单纯的“季节”变动对事件影响的大小。比如 3 月份的季节指数为 0.551，而 6 月份的季节指数为 1.919，这说明 6 月份的平均气温一般是 3 月份的 3.48 倍，值如明年 3 月份北京市的平均气温为 7℃，那么根据季节效应我们有理由预测在大气环境不发生大的改变的前提下，

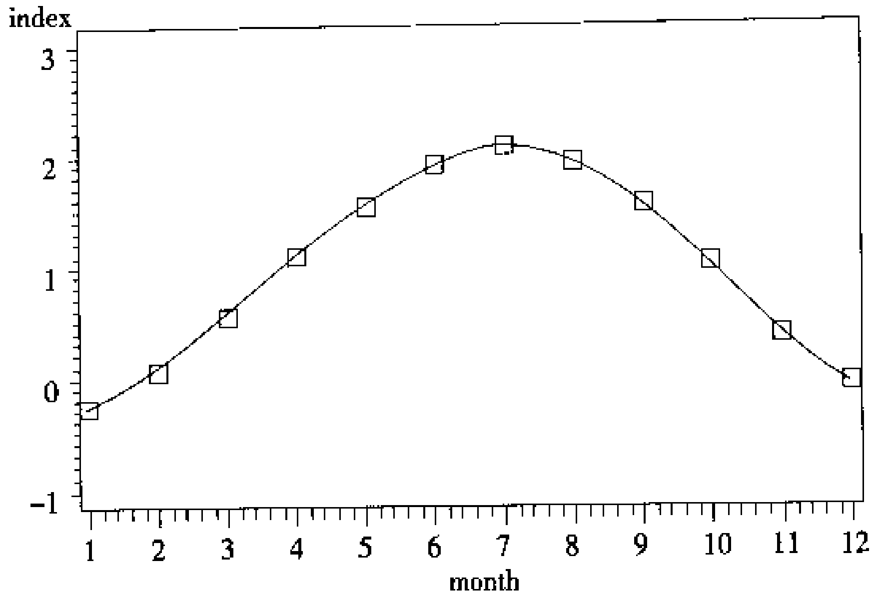


图 4—5 北京市 1995—2000 年月平均气温序列季节指数图

北京市 6 月份的平均气温应该在 21℃左右。

4.5 综合分析

在上面两节里我们分别介绍了单纯的趋势分析方法和单纯的季节效应分析方法。在这一节里我们主要介绍既有趋势起伏变动又有季节效应的复杂序列的分析方法。

要对趋势起伏变动和季节效应同时进行分析，首先必须了解它们之间的相互作用关系。但是，通常我们并不知道它们到底是怎样综合作用而影响序列变化的，这时只能根据序列的表现，选择一些经验模型来估计各因素之间的相互作用关系。常用的模型有：

（1）加法模型

$$x_t = T_t + S_t + I_t \tag{4.3}$$

（2）乘积模型

$$x_t = T_t \cdot S_t \cdot I_t \tag{4.4}$$

（3）混合模型

(a) $x_t = S_t \cdot T_t + I_t$

(b) $x_t = S_t \cdot (T_t + I_t)$ （4.5）

式中，T_t 代表序列的长期趋势波动；S_t 代表序列的季节性（周期性）变化；I_t

代表随机波动。

在确定性影响很强劲而不确定性影响很微弱时，选择合适的确定性模型通常会得到非常不错的分析预测结果。

【例 4.7】 对 1993—2000 年中国社会消费品零售总额序列（数据见附录 1.11）进行确定性时序分析。

（1）绘制时序图。如图 4—6 所示。

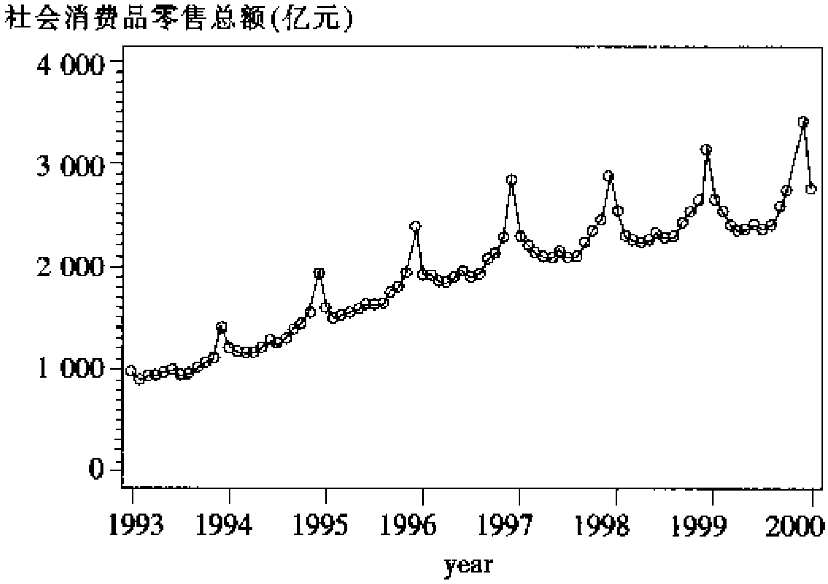


图 4—6 中国社会消费品零售总额时序图

（2）选择拟合模型。从该序列的时序图中，我们看到长期递增趋势和以年为固定周期的季节波动同时作用于该序列，因而尝试使用混合模型（4.5b）拟合该序列的发展：

$$x_t = S_t \cdot (T_t + I_t)$$

（3）计算该序列的季节指数 $\hat{S}_i$（$i=1, 2, \cdots, 12$）。根据附录 1.11 的数据，容易算出该序列的月度季节指数如表 4—5 所示。

表 4—5

月份	季节指数	月份	季节指数
1	0.982	7	0.929
2	0.943	8	0.940
3	0.920	9	1.001
4	0.911	10	1.054
5	0.925	11	1.100
6	0.951	12	1.335

绘制季节指数图，如图 4—7 所示。

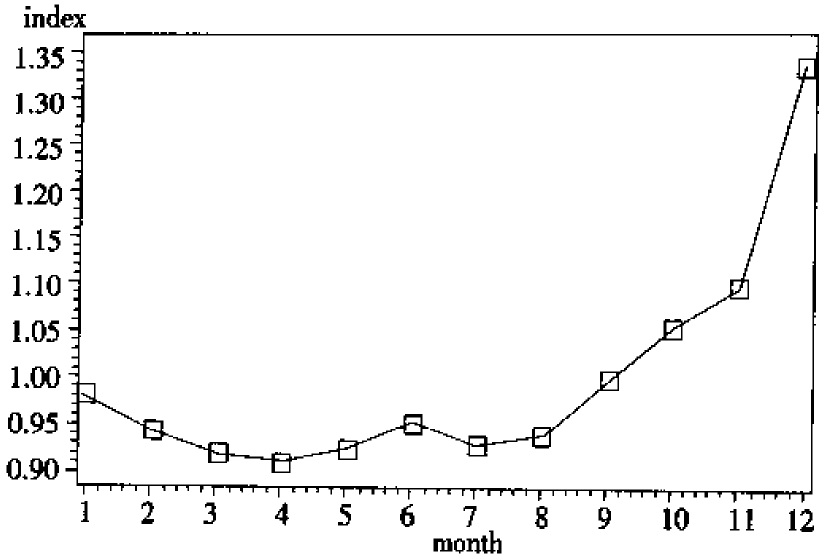


图 4—7　中国社会消费品零售总额序列季节指数图

从季节指数图可以非常直观地看出每年的四季度是我国社会消费品销售旺季，而前三个季度的销售状况起伏不大，季节数应不明显。

（4）消除季节影响后拟合该序列的趋势 $\hat{T}_t$ 变动规律。根据拟合模型假设 $x_t = S_t \cdot (T_t + I_t)$，原始序列值除以相应的季节指数，就基本上消除了季节性因素对原序列的影响，而只剩下长期趋势波动和随机波动的影响：

$$\frac{x_t}{\hat{S}_t} = T_t + I_t$$

图 4—8 显示该序列有一个基本线性递增的长期趋势。于是考虑用一元线性回归进行趋势拟合：

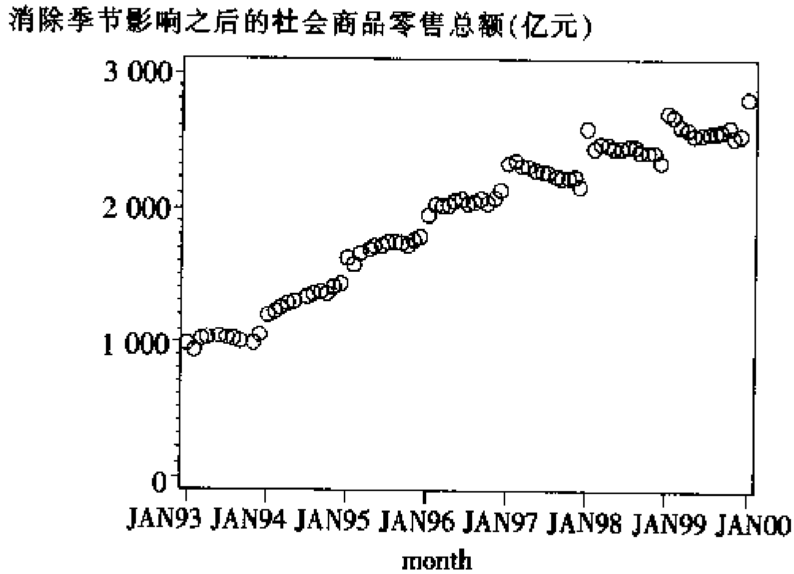


图 4—8　消除季节影响之后的社会商品零售总额序列散点图

$$\hat{T}_t = a + bt$$

用最小二乘估计方法，容易得到该线性趋势模型为：

$$\hat{T}_t = 1\,015.522 + 20.931\,78t,\ t = 1,\ 2,\ \cdots,\ 96$$

线性趋势拟合后的效果图如图 4—9 所示。

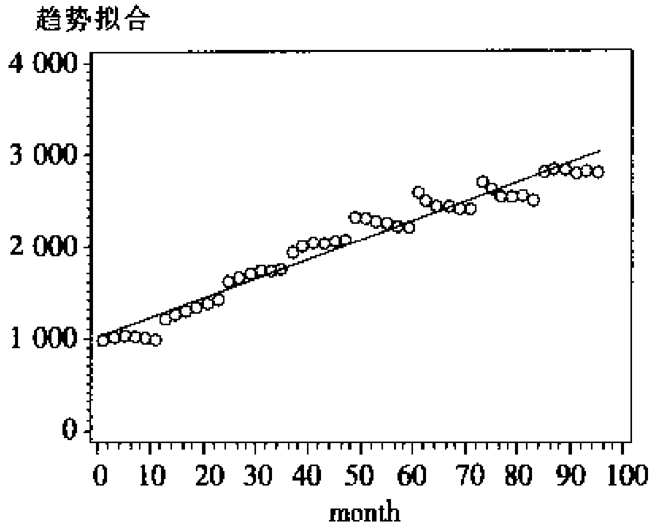


图 4—9　线性趋势拟合图

（5）残差检验。用原序列值除以季节指数，再减去长期趋势拟合值之后的残差项就可以视为随机波动的影响。

$$\frac{x_i}{\hat{S}_t} - \hat{T}_t = I_t$$

本例残差图如图 4—10 所示。

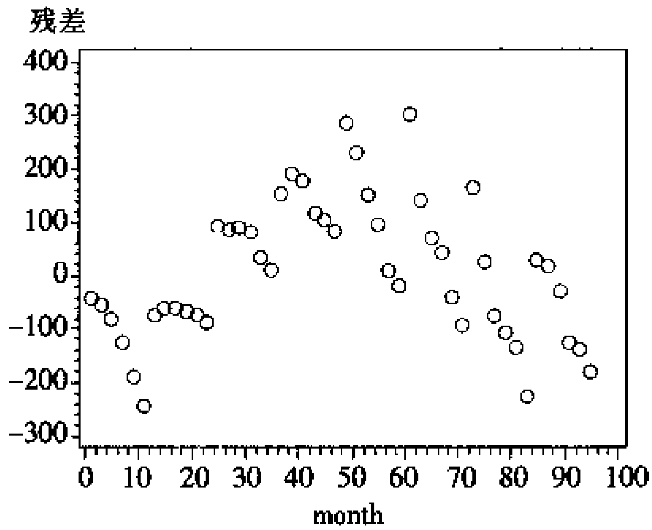


图 4—10　残差图

残差图显示残差序列仍然存在一定的相关性。这说明我们拟合的这个模型还没有把原序列中蕴含的相关信息充分提取出来，这是确定性分析方法常见的缺

点。因素分解方法的侧重点在于确定性信息快速、便捷地提取，但对于信息提取的充分性常常不能达到完美。

（6）短期预测。利用拟合模型可以对序列进行短期预测，第 t 期的预测值为：

$$\hat{x}_t(l)=\hat{S}_{t-l}\cdot\hat{T}_{t+l}$$

利用预测模型和历史数据，得到 2001 年各月份中国社会消费品零售总额季节指数、趋势值及预测值如表 4—6 所示。

表 4—6

预测期数	$\hat{S}_{96-l}$	$\hat{T}_{96+l}$	$\hat{x}_{96}(l)$
1	0.982 155	3 045.905	2 991.551
2	0.943 092	3 066.836	2 892.308
3	0.919 868	3 087.768	2 840.339
4	0.910 837	3 108.700	2 831.519
5	0.925 414	3 129.632	2 896.205
6	0.951 451	3 150.564	2 997.607
7	0.929 076	3 171.495	2 946.560
8	0.939 645	3 192.427	2 999.748
9	1.009 497	3 213.359	3 243.876
10	1.053 718	3 234.291	3 408.031
11	1.100 457	3 255.222	3 582.232
12	1.334 789	3 276.154	4 372.974

将 1993—2000 年中国社会消费品零售总额观察值与估计值序列联合作图，如图 4—11 所示。

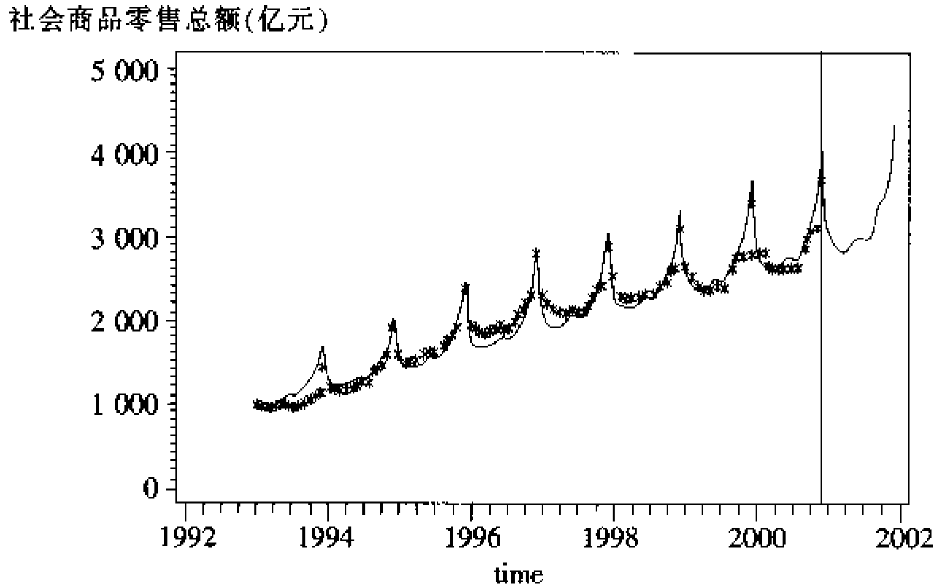


图 4—11　拟合效果图

图中星号表示观察值数据，曲线表示预测值时序图。根据拟合图的直观显示，可以看出我们所拟合的确定性时序分析模型对该序列总体变化规律的把握还是比较准确的。这说明在本例中，尽管确定性因素分解方法对信息的提取不够充分，但是由于确定性因素影响非常强劲，相比而言，残差序列的影响非常微弱，遗漏残差序列中蕴含的小部分相关信息，对模型的拟合精度没有显著性的影响。所以该确定性因素分解模型仍然是显著有效的。

4.6 X-11 过程

X-11 过程是美国国情调查局编制的时间序列季节调整过程。它的基本原理就是时间序列的确定性因素分解方法。

X-11 过程基于这样的假定：任何时间序列都可以拆分成长期趋势起伏 T_t、季节波动 S_t、不规则波动 I_t 的影响。又有经济学家发现在经济时间序列中交易日 D_t 也是一个很重要的影响因素，因此任何一个时间序列可以如下分解：

乘法模型：$x_t=T_tS_tD_tI_t$

加法模型：$x_t=T_t+S_t+D_t+I_t$

宏观调控部门主要关注的是长期趋势起伏项 T_t 的规律，所以 X-11 过程主要是要从原序列中剔除季节影响、交易日影响及随机影响，得到尽量准确的长期趋势规律。而实现这个目标的方法就是前面分绍的因素剔除法。

在具体的操作方法上它不依赖任何模型，这是因为：(1) 季节调整方法始于 20 世纪二三十年代，那时还不存在合适的分析模型；(2) 在历史数据的基础上拟合的模型只能有效地表达一段有限长度的时间序列的发展变化规律，一旦加入更多数据，模型就不具有很好的解释性。由于这两个原因，所以在 X-11 过程中普遍采用移动平均的方法：用多次短期中心移动平均消除随机波动，用周期移动平均消除趋势，用交易周期移动平均消除交易日影响。在整个过程中总共要用到 11 次移动平均，所以得名 X-11 过程。

【例 4.7 续】 对 1993—2000 年中国社会消费品零售总额序列使用 X-11 过程进行季节调整。

由于没有交易日的影响，我们考虑乘法模型

$$x_t=T_tS_tI_t$$

使用 X-11 过程得到平均季节指数如表 4—7 所示。

表 4—7

月份	1	2	3	4	5	6
季节指数	1.046	0.995	0.959	0.939	0.943	0.960
月份	7	8	9	10	11	12
季节指数	0.925	0.924	0.981	1.010	1.052	1.267

绘制相应的季节指数图如图 4—12 所示。

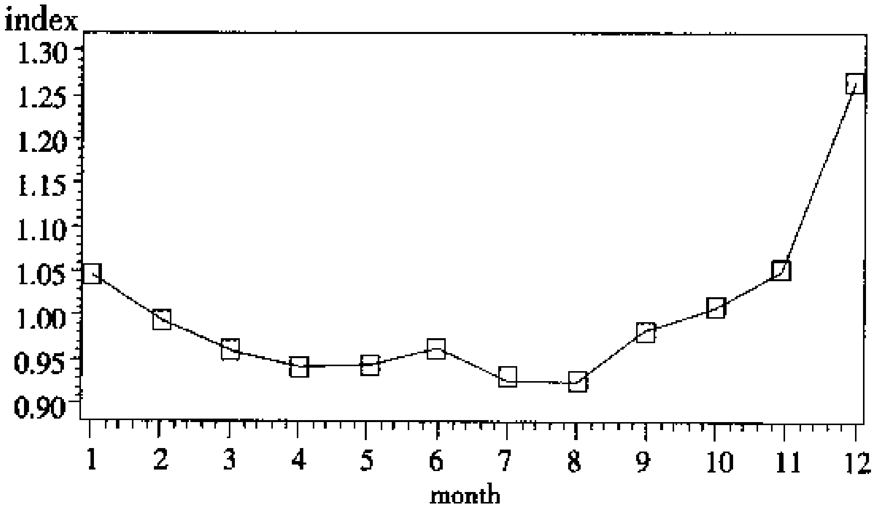


图 4—12　X-11 过程获得的平均季节指数图

和图 4—7 对比，可以发现这两个季节指数图形状基本一致，但数值不完全相同。通常 X-11 过程估计的季节指数更准确一些。这是因为在有趋势时使用例 4.7 所述的因素分解方法对季节指数的估计通常是有偏的，同时，在季节效应存在时使用例 4.7 的方法对趋势项的估计也是不准确的。而 X-11 过程的多次移动平均并使用了迭代方法，有效地提高了对季节指数及趋势拟合的精度。

消除季节趋势，得到的调整后序列图如图 4—13 所示。

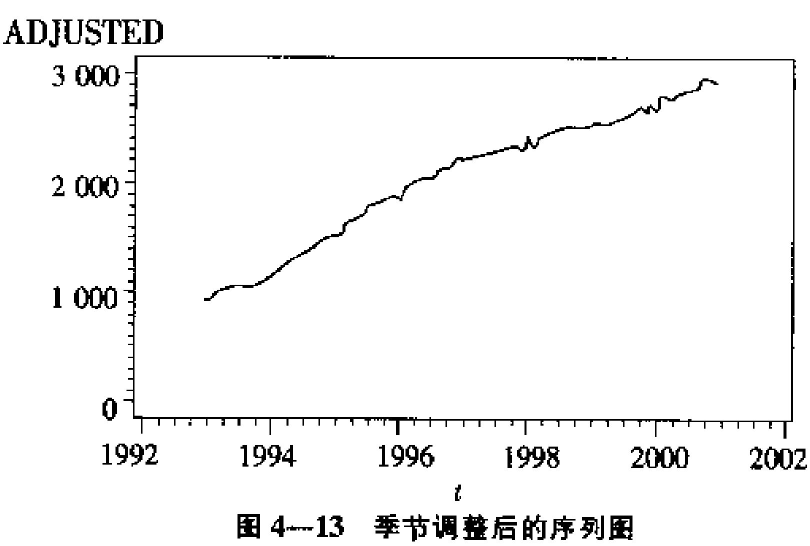


图 4—13　季节调整后的序列图

可以看出中国社会消费品零售总额剔除季节效应之后有非常显著的线性递增趋势。这和例 4.7 得到的结论一致。使用移动平均的方法，拟合序列的趋势，所得趋势拟合图如图 4—14 所示。

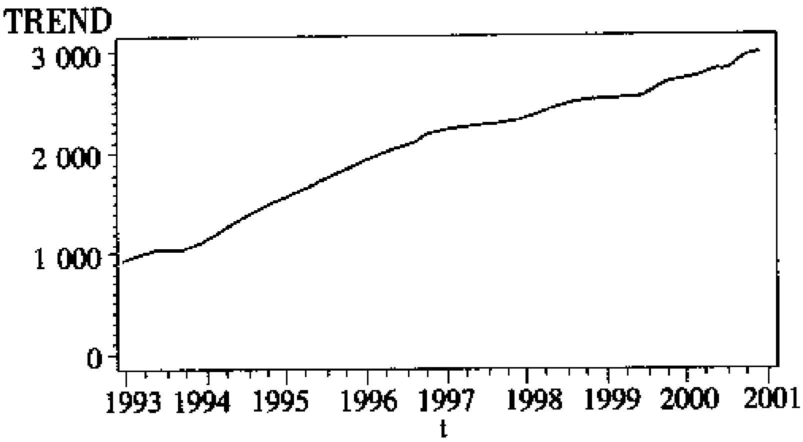


图 4—14　季节调整后的趋势拟合图

从季节调整后序列中消除趋势项，得到随机波动项，见图 4—15。

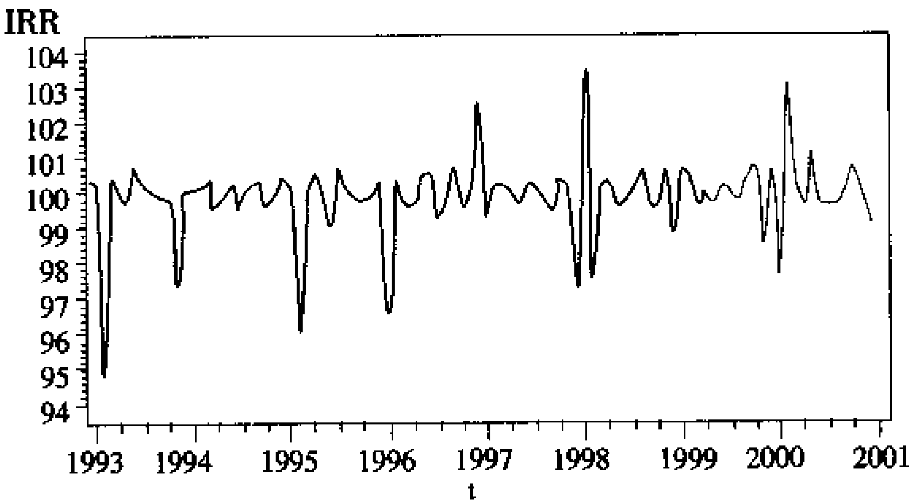


图 4—15　随机波动项时序图

对照例 4.7 的残差图（图 4—10），可以直观看出 X-11 过程得到的残差序列（图 4—15）比简单的因素分解方法得到的残差序列（图 4—10）更不规则。这说明 X-11 过程对季节效应和趋势信息的提取更加充分。

4.7 习　题

1. 使用 4 期移动平均作预测，求在 2 期预测使 $\hat{x}_{T+2}$ 中 x_T 与 x_{T-1} 前面的系数

分别等于多少?

2. 使用指数平滑法得到 $\tilde{x}_{t-1}=5$，$\tilde{x}_{t+1}=5.26$，已知序列观察值 $x_t=5.25$，$x_{t+1}=5.5$，求指数平滑系数 α。

3. 有一 20 期的观察值序列 $\{x_t\}$ 为：

10 11 12 10 11 14 12 13 11 15 12 14 13 12 14 12 10 10 11 13

(1) 使用 5 期移动平均法预测 $\hat{x}_{22}$。

(2) 使用指数平滑法确定 $\hat{x}_{22}$，其中平滑系数为 $\alpha=0.4$。

(3) 假设 a 为 5 期移动平均法预测 $\hat{x}_{22}$ 中 x_{20} 前的系数，b 为平滑系数为 $\alpha=0.4$ 的指数平滑法预测 $\hat{x}_{22}$ 中 x_{20} 前的系数，求 $b-a$。

4. 现有序列 $\{x_t=t,\ t=1,\ 2,\ \cdots\}$ 使用平滑系数为 α 的指数平滑法修匀该序列。假定 $\tilde{x}_0=0$，求 $\lim\limits_{t\to\infty}\frac{\tilde{x}_t}{t}$。

5. 对某观察值序列 $\{x_t\}$ 使用 Holt 两参数指数平滑方法，其中两个平滑系数分别为 $\alpha=0.4$，$\gamma=0.2$。假定已知 $\tilde{x}_{t-1}=20$，$r_{t-1}=5$，$r_t=4.1$，求 x_t。

6. 爱荷华州 1948—1979 年非农产品季度收入数据如表 4—8 所示。

表 4—8

601	604	620	626	641	642	645	655	682	678	692	707
736	753	763	775	775	783	794	813	823	826	829	831
830	838	854	872	882	903	919	937	927	962	975	995
1 001	1 013	1 021	1 028	1 027	1 048	1 070	1 095	1 113	1 143	1 154	1 173
1 178	1 183	1 205	1 208	1 209	1 223	1 238	1 245	1 258	1 278	1 294	1 314
1 323	1 336	1 355	1 377	1 416	1 430	1 455	1 480	1 514	1 545	1 589	1 634
1 669	1 715	1 760	1 812	1 809	1 828	1 871	1 892	1 946	1 983	2 013	2 045
2 048	2 097	2 140	2 171	2 208	2 272	2 311	2 349	2 362	2 442	2 479	2 528
2 571	2 634	2 684	2 790	2 890	2 964	3 085	3 159	3 237	3 358	3 489	3 588
3 624	3 719	3 821	3 934	4 028	4 129	4 205	4 349	4 463	4 598	4 725	4 827
4 939	5 067	5 231	5 408	5 492	5 653	5 828	5 965				

选择适当模型拟合该序列长期趋势。

7. 某地区 1962--1970 年平均每头奶牛的月度产奶量数据（单位：磅）如表 4—9 所示。

表 4—9

589	561	640	656	727	697	640	599	568	577	553	582
600	566	653	673	742	716	660	617	583	587	565	598
628	618	688	705	770	736	678	639	604	611	594	634
658	622	709	722	782	756	702	653	615	621	602	635
677	635	736	755	811	798	735	697	661	667	645	688

续前表

713	667	762	784	837	817	767	722	681	687	660	698
717	696	775	796	858	826	783	740	701	706	677	711
734	690	785	805	871	845	801	764	725	723	690	734
750	707	807	824	886	859	819	783	740	747	711	751

(1) 绘制该序列时序图，直观考察该序列的特点。

(2) 使用因素分解方法，拟合该序列的发展，并预测 1976 年该地区奶牛的月度产量。

(3) 使用 X-11 方法，确定该序列的趋势。

8. 某城市 1980 年 1 月至 1995 年 8 月每月屠宰生猪数量（单位：头）如表 4 −10 所示。

表 4—10

76 378	71 947	33 873	96 428	105 084	95 741	110 647	100 331	94 133	103 055
90 595	101 457	76 889	81 291	91 643	96 228	102 736	100 264	103 491	97 027
95 240	91 680	101 259	109 564	76 892	85 773	95 210	93 771	98 202	97 906
100 306	94 089	102 680	77 919	93 561	117 062	81 225	88 357	106 175	91 922
104 114	109 959	97 880	105 386	96 479	97 580	109 490	110 191	90 974	98 981
107 188	94 177	115 097	113 696	114 532	120 110	93 607	110 925	103 312	120 184
103 069	103 351	111 331	106 161	111 590	99 447	101 987	85 333	86 970	100 561
89 543	89 265	82 719	79 498	74846	73 819	77 029	78 446	86 978	75 878
69 571	75 722	64 182	77 357	63292	59 380	78 332	72 381	55 971	69 750
85 472	70 133	79 125	85 805	81778	86 852	69 069	79 556	88 174	66 698
72 258	73 445	76 131	86 082	75443	73 969	78 139	78 646	66 269	73 776
80 034	70 694	81 823	75 640	75540	82 229	75 345	77 034	78 589	79 769
75 982	78 074	77 588	84 100	97966	89 051	93 503	84 747	74 531	91900
81 635	89 797	81 022	78 265	77271	85 043	95 418	79 568	103 283	95 770
91 297	101 244	114 525	101 139	93 866	95 171	100 183	103 926	102 643	108 387
97 077	90 901	90 336	88 732	83 759	99 267	73 292	78 943	94 399	92 937
90 130	91 055	106 062	103 560	104 075	101 783	93 791	102 313	82 413	83 534
109 011	96 499	102 430	103 002	91 815	99 067	110 067	101 599	97 646	104 930
88 905	89 936	106 723	84 307	114 896	106 749	87 892	100 506		

选择适当模型拟合该序列的发展，并预测 1995 年 9 月至 1997 年 9 月该城市生猪屠宰数量。

4.8　上机指导

4.8.1　拟合线性趋势

在 SAS 系统中 REG（回归）过程与 AUTOREG（自回归）过程都可以进行

时间序列线性趋势拟合。假定我们要分析的数据存于临时数据集 example4 _ 1 中，要拟合的线性回归模型为 $x_t=a+bt$，则相关命令如下：

```
data example4_1;
input x@@;
t=_n_;
cards;
  12.79  14.02  12.92  18.27  21.22  18.81
  25.73  26.27  26.75  28.73  31.71  33.95
;
proc autoreg data=example4_1;
model x=t;
run;
```

运行该程序输出结果如图 4—16 所示。

```
                     The AUTOREG Procedure
                     Dependent Var iable X
                 Ordinary Least Squares Estimates
SSE                    26.190 220 6   DFE                       10
MSE                     2.619 02      Root MSE             1.618 34
SBC                    48.390 091 3   AIC                47.420 278
Regress R-Square        0.955 5       Total R-Square         0.955 5
Durbin-Watson           2.726 9

                                            Standard             Approx
  Var iable      DF      Estimate       Error      t Value     Pr> |t|
Intercept         1        9.708 6      0.996 0      9.75       <.000 1
t                 1        1.982 9      0.135 3     14.65       <.000 1
```

图 4—16　AUTOREG 过程输出线性拟合结果

该输出窗口共输出三方面的信息。

1. 因变量的名称，本例中因变量为 x。

2. 普通最小二乘估计相关统计量，该部分输出信息如表 4—11 所示。

表 4—11

SSE（误差平方和）	DFE（误差平方和的自由度）
MSE（均方误差）	Root MSE（均方根误差）
SBC（SBC 信息量）	AIC（AIC 信息量）
Regress R-Square（只针对回归模型的 r^2）	Total-R-Square（包括自回归误差过程在内的整体模型的 r^2）
Durbin-Watson（DW 统计量）	

3. 参数估计值。该部分从左到右输出的信息分别是：变量名、自由度、估计值、估计值的标准差、t 值以及 t 值的近似 P 值。

4.8.2 拟合非线性趋势

在 SAS 系统中有一个 NLIN（非线性）过程可以进行时间序列非线性趋势拟合。

假定我们要分析的数据存于临时数据集 example4 _ 2 中，要拟合的非线性回归模型为 $x_t = at + b^t$，则相关命令如下：

```
data example4_2;
input x@@;
t=_n_;
cards;
   1.85     7.48     14.29     23.02     37.42     74.27    140.72
 265.81   528.23  1 040.27  2 064.25  4 113.73  8 212.21 16 405.95
;
proc nlin method=gauss;
model x=a*t+b**t;
parameters a=0.1 b=0.1;
der.a=t;
der.b=log(b)*b**t;
output predicted=xhat out=out;
run;
```

语句说明：

1. "proc nlin method=gauss;" 指令系统采用 GAUSS 迭代法进行非线性参数估计。在 NLIN 过程中一共允许选择五种迭代方法，它们分别是

NEWTON——牛顿迭代法。

GAUSS——高斯迭代法，也称为改良的高斯一牛顿法。

MARQUARDT——马克特迭代法。

GRANDIENT——梯度法，也称为最速下降法（steepest descent method）。

DUD——又称为错位法（false position）、多元割线法（multivariate secant）。

其中前三种迭代法的迭代功能强于后两种方法。如果不特别指定迭代方法，系统默认使用高斯迭代法。

2. "model x=a*t+b**t;" 告诉系统拟合模型结构。

3. "parameters a=0.1 b=0.1;" 告诉系统待估参数是哪些，并给出待估参数的迭代初始值。

4. “der. a=t; der. b=log(b) * b * * t;” 给出待估参数的一阶导函数，以便于迭代计算。

5. “output predicted=xhat out=out; ” 输出部分结果到临时数据集 OUT，输出内容时间 t，原序列值 x_t 和估计值 $\hat{x}_t$（本例中该变量取名为 XHAT）。

运行该程序共输出如下六方面信息：

1. 迭代过程。如图 4—17 所示。

Iterative Phase			
Iter	a	b	Sum of Squares
0	0. 100 0	0. 100 0	3. 591 5E8
1	0. 305 9	1. 187 8	3. 583 5E8
2	0. 238 2	1. 954 5	26 413 326
3	0. 215 2	1. 956 0	24 954 991
4	0. 099 6	1. 963 4	18 068 239
5	−0. 478 7	2. 001 7	36 024. 6
6	−0. 178 0	1. 999 5	11 127. 6
7	0. 107 8	2. 000 4	1 420. 5
8	0. 193 2	2. 000 2	1 140. 7

图 4—17　NLIN 过程输出迭代结果

2. 收敛状况。如图 4—18 所示。

WARN ING: PROC NL IN failed to converge.

图 4—18　NLIN 过程输出迭代结果

这是警告我们本次迭代没有收敛。

3. 估计信息摘要。如图 4—19 所示。

Estimation Summary (Not Converged)	
Method	Gauss-Newton
Iterations	8
Subiterations	55
Average Subiterations	6. 875
R	0. 987 047
PPC (a)	7. 254 293
PPC	.
Object	0. 196 937
Objective	1 140. 726
Observations Read	14
Observations Used	14
Observations Missing	0

图 4—19　NLIN 过程输出估计信息

4. 主要统计量。如图 4—20 所示。

Source	DF	Sum of Squares	Mean Square	F Value	Approx Pr> F
Regression	2	3.592 4E8	1.796 2E8	1 889 521	<.000 1
Residual	12	1 140.7	95.060 5		
Uncorrected Total	14	3.592 4E8	95.060 5		
Corrected Total	13	2.817 9E8			

图 4—20　NLIN 过程输出主要统计量信息

5. 参数信息摘要。如图 4—21 所示。

		Approx		
Parameter	Estimate	Std Error	Approximate 95%	Confidence Limits
a	0.193 2	0.432 6	−0.749 3	1.135 7
b	2.000 2	0.001 05	1.997 9	2.002 4

图 4—21　NLIN 过程输出迭代参数估计信息

本例中，得到的拟合模型为：

$$x_t = 0.193\,2t + 2.002^t + \varepsilon_t$$

6. 近似相关阵。如图 4—22 所示。

Approximate Correlation Matrix		
	a	b
a	1.000 000 0	−0.706 733 7
b	−0.706 733 7	1.000 000 0

图 4—22　NLIN 过程输出参数估计值的近似相关阵

为了直观地看出拟合效果，我们可以将原序列值和拟合值联合作图，相关命令如下：

```
proc gplot data=out;
plot x * t=1 xhat * t=2/overlay;
symbol1 c=black i=none v=star;
symbol2 c=red i=join v=none;
run;
```

输出图像如图 4—23 所示。

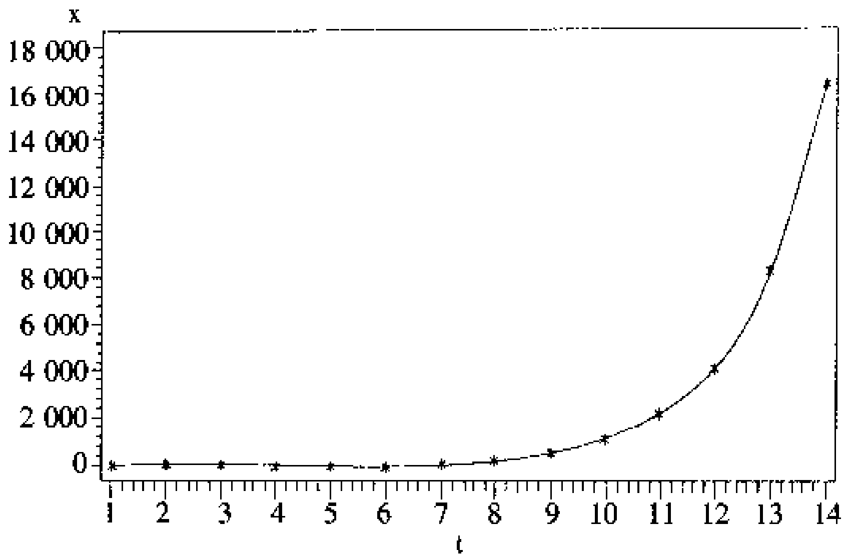


图 4—23 NLIN 过程拟合效果图

图中星号为原序列观察值，曲线为拟合值。通过该图可以看出尽管迭代过程最终并没有收敛，但拟合效果还是非常不错的。

4.8.3 X-11 过程

以 1978—1988 年英国非耐用消费额为例，将数据读入临时数据集 example4_3，使用 X-11 过程进行季节调整，并将原序列与消除季节影响的趋势线联合做图输出，相关命令如下：

```
data example4_3;
input x@@;
t=intnx('quarter','1jan1978'd, _n_-1);
format t yyq4.;
cards;
40777  41778  43160  45897
41947  44061  44378  47237
43315  43396  44843  46835
42833  43548  44637  47107
42552  43526  45039  47940
43740  45007  46667  49325
44878  46234  47055  50318
46354  47260  48883  52605
```

```
48527   50237   51592   55152
50451   52294   54633   58802
53990   55477   57850   61978
;
proc x11 data=example4_3;
quarterly date=t;
var x;
output out=out b1=x d10=season d11=adjusted d12=trend d13=irr;
proc gplot data=out;
plot x*t=1 adjusted*t=2/overlay;
plot adjusted *t=1
trend *t=1
irr*t=1;
symbol1 c=black i=join v=star;
symbol2 c=red i=join v=none w=2 l=3;
run;
```

语句说明：

(1) " proc x11 data=example4_3;" 指令系统对数据集 example4_3 的数据进行 X-11 分析。

(2) "quarterly date=t;" 告诉系统这是季度数据（假如是月度数据就应该记作 monthly），变量 t 为时间变量名。

(3) "var x;" 告诉系统要进行季节调整的变量为 x。

(4) "output out=out b1=x d10=season d11=adjusted d12=trend d13=irr;" 告诉系统输出部分结果到临时数据集 OUT，在此我们要求输出的结果是：

原序列值 x（表 b1 的数值）；

季节指数（或称为季节因子）season（表 d10 的数据）；

季节调整后的序列值 adjusted（表 d11 的数据）；

趋势拟合值 trend（表 d12 的数据）；

及最后的不规则波动值 irr（表 d13 的数据）。

X-11 过程的输出结果非常多，主要分为七个部分：

A. 先验修正（可选）

B. 不规则成分权重和回归交易日因子的初始估计

C. 规则成分权重和回归交易日因子的最终估计

D. 季节项、长期趋势项和不规则波动的最终估计

E. 分析表格

F. 概括性量度

G. 图表

在每一部分输出里又包含了非常多的子表，由于内容过于庞杂，详细的解释略。每张表的具体输出内容及解释请查考附录 3。

本例输出的季节指数图、消除季节效应后的序列图、趋势图及不规则波动图略，输出的原序列与消除季节效应后的调整序列图如图 4—24 所示。

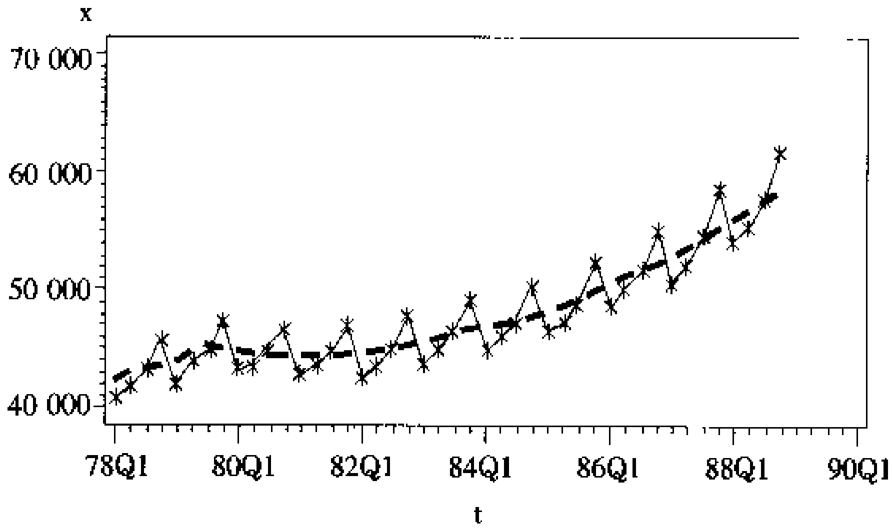


图 4—24　X-11 过程季节调整前后的效果图

第 5 章

非平稳序列的随机分析

第 4 章介绍了非平稳序列的确定性因素分解方法。这种方法具有原理简单、操作简便、易于解释等优点，在宏观经济管理与预测领域有着广泛的应用。

但是随着研究方法的深入和研究领域的拓宽，人们发现确定性因素分解方法还存在一些问题，主要有：

1. 确定性因素分解方法只能提取强劲的确定性信息，对随机性信息浪费严重。

2. 确定性因素分解方法把所有序列的变化都归结为四大因素的综合影响，却始终无法提供明确、有效的方法判断各大因素之间确切的作用关系。

这些问题导致确定性因素分解方法不能充分提取观察值序列中的有效信息，导致模型拟合精度通常不够理想。随机时序分析方法的发展就是为了弥补确定性因素分解方法的不足，为人们提供了更加丰富、更加精确的时序分析工具。

5.1 差分运算

5.1.1 差分运算的实质

拿到观察值序列之后，无论是采用确定性时序分析方法还是随机时序分析方法，分析的第一步都是要通过有效的手段提取序列中所蕴含的确定性信息。

确定性信息的提取方法非常多，前面我们介绍过的构造季节指数，拟合长期趋势模型、移动平均、指数平滑等诸多方法都是确定性信息提取方法。但是它们对确定性信息的提取都不够充分。

Cox 和 Jenkins 在 *Time Series Analysis Forecasting and Control* 一书中特别强调差分方法的使用，他们使用大量的案例分析证明差分方法是一种非常简便、有效的确定性信息提取方法。而 Cramer 分解定理则在理论上保证了适当阶数的差分一定可以充分提取确定性信息。

根据 Cramer 分解定理，方差齐性非平稳序列都可以分解为如下形式：

$$x_t = \sum_{j=1}^{d}\beta_j \cdot t^j + \Psi(B)a_t$$

式中，$\{a_t\}$ 为零均值白噪声序列。

离散序列的 d 阶差分就相当于连续变量的 d 阶求导，显然，在 Cramer 分解定理的保证下，d 阶差分就可以将 $\{x_t\}$ 中蕴含的确定性信息充分提取。

$$\nabla^d \sum_{j=1}^{d}\beta_j t^j = c, c \text{ 为某一常数}$$

展开 1 阶差分，有

$$\nabla x_t = x_t - x_{t-1}$$

等价于

$$x_t = x_{t-1} + \nabla x_t$$

这意味着 1 阶差分实质上就是一个自回归过程，它是用延迟一期的历史数据 $\{x_{t-1}\}$ 作为自变量来解释当期序列值 $\{x_t\}$ 的变动状况，差分序列 $\{\nabla x_t\}$ 度量的是 $\{x_t\}$ 1 阶自回归过程中产生的随机误差的大小。

展开任意一个 d 阶差分，有

$$\nabla^d x_t = (1-B)^d x_t = \sum_{i=0}^{d}(-1)^i C_d^i x_{t-i}$$

它的实质就是一个 d 阶自回归过程。

$$x_t = \sum_{i=1}^{d}(-1)^{i+1}C_d^i x_{t-i} + \nabla^d x_t$$

这意味着差分运算的实质是使用自回归的方式提取确定性信息。

5.1.2 差分方式的选择

实践中，我们会根据序列的不同特点选择合适的差分方式，常见情况有如下三种：

一、序列蕴含着显著的线性趋势，1 阶差分就可以实现趋势平稳

【例 5.1】 在例 2.1 的分析中，我们知道 1964—1999 年中国纱年产量序列蕴含着一个近似线性的递增趋势。现对该序列进行 1 阶差分运算，考察差分运算对该序列线性趋势信息的提取作用。

利用附录 1.2 所提供的数据，对原序列进行 1 阶差分运算：

$$\nabla x_t = x_t - x_{t-1}$$

差分后序列 $\{\nabla x_t\}$ 时序图如图 5—1 所示。

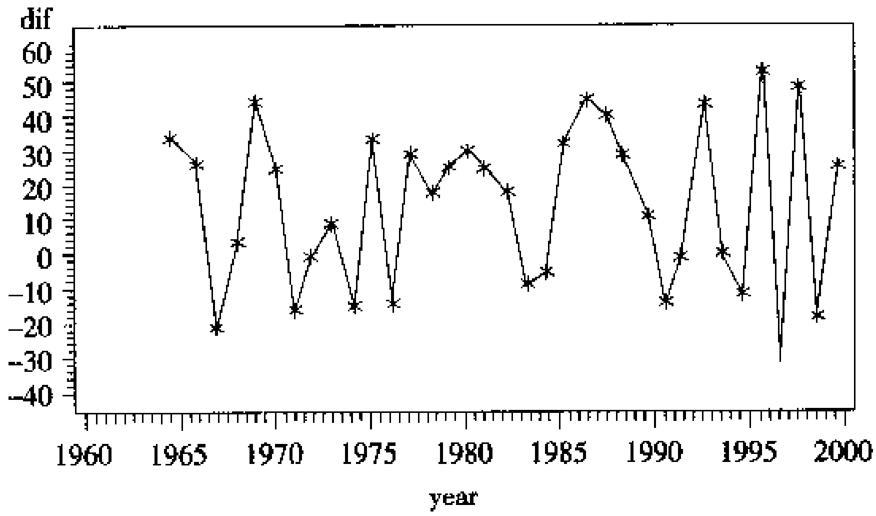


图 5—1 1964—1999 年中国纱年产量 1 阶差分序列时序图

时序图清晰地显示，1 阶差分运算已经非常成功地从原序列中提取出了线性趋势，差分后序列呈现出非常平稳的随机波动。

二、序列蕴含着曲线趋势，通常低阶（2 阶或 3 阶）差分就可以提取出曲线趋势的影响

【例 5.2】 尝试提取 1950—1999 年北京市民用车辆拥有量序列（数据见附录 1.12）的确定性信息。

该序列时序图如图 5—2 所示。

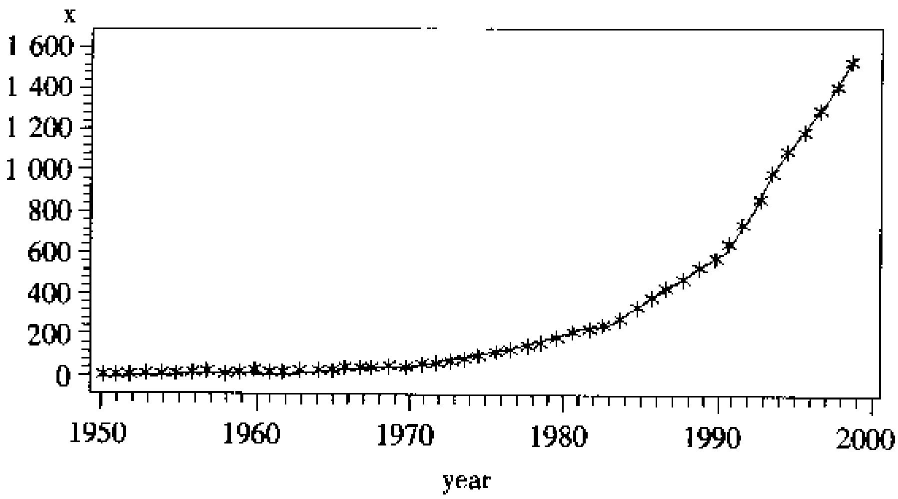


图 5—2　1950—1999 年北京市民用车辆拥有量时序图

从时序图中我们可以清楚地看到该序列蕴含着曲线递增的长期趋势。对该序列进行 1 阶差分运算：

$$\nabla x_t = x_t - x_{t-1}$$

1 阶差分后序列 $\{\nabla x_t\}$ 时序图如图 5—3 所示。

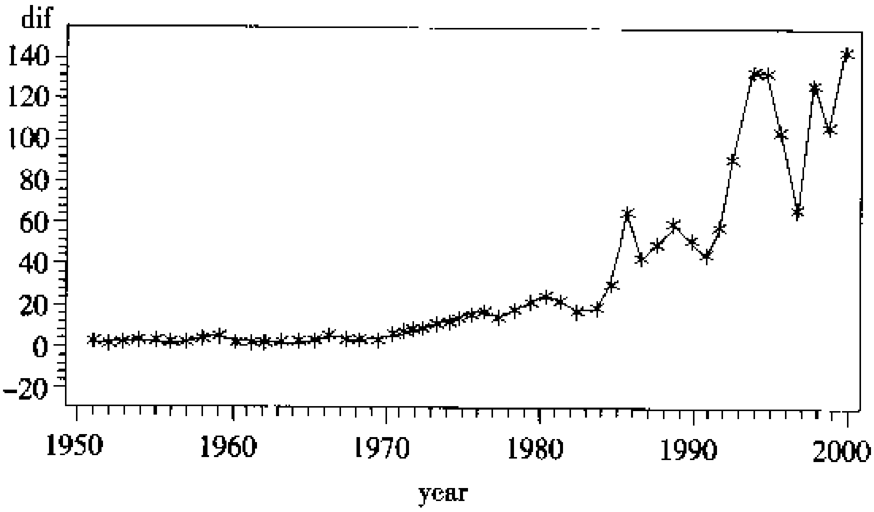


图 5—3　1950—1999 年北京市民用车辆拥有量 1 阶差分后时序图

图 5—3 显示，1 阶差分提取了原序列中部分长期趋势，但长期趋势信息提取不充分，1 阶差分后序列中仍蕴含着长期递增的趋势，于是对 1 阶差分后序列再做一次差分运算：

$$\nabla^2 x_t = \nabla x_t - \nabla x_{t-1}$$

2 阶差分后序列 $\{\nabla^2 x_t\}$ 时序图如图 5—4 所示。

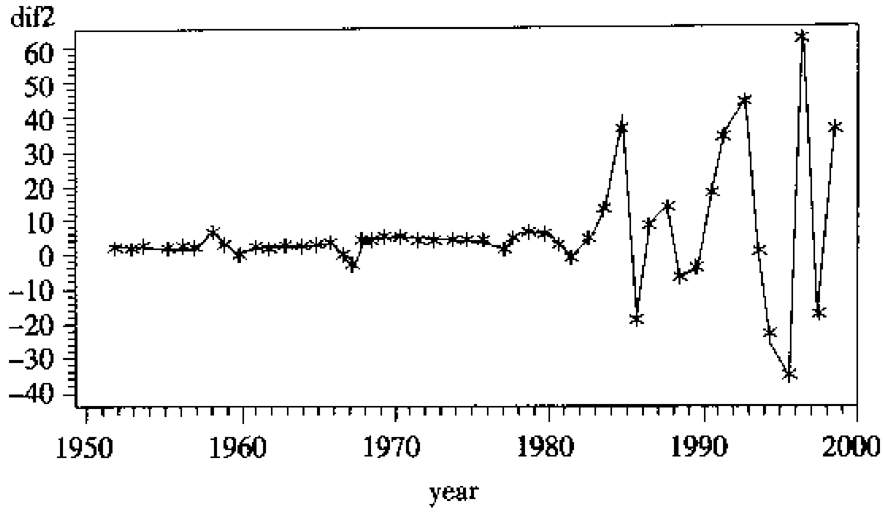


图 5—4　1950—1999 年北京市民用车辆拥有量 2 阶差分后时序图

图 5—4 显示，2 阶差分比较充分地提取了原序列中蕴含的长期趋势，使得差分后序列不再呈现确定性趋势了。

三、蕴含固定周期的序列

对于蕴含着固定周期的序列进行步长为周期长度的差分运算，通常可以较好地提取周期信息。

【例 5.3】　差分运算提取 1962 年 1 月—1975 年 12 月平均每头奶牛的月产奶量序列中的确定性信息（数据见附录 1.13）。

该序列时序图如图 5—5 所示。

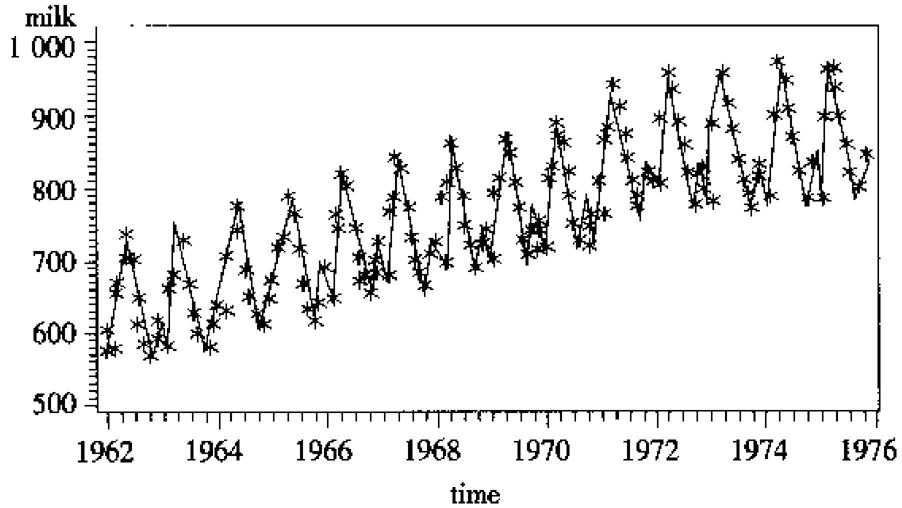


图 5—5　奶牛的月产奶量序列时序图

时序图显示该序列具有一个线性递增的长期趋势和一个周期长度为一年的稳

定的季节变动。所以对原序列先做一步差分，提取线性递增趋势：

$$\nabla x_t = x_t - x_{t-1}$$

一步差分后序列 $\{\nabla x_t\}$ 时序图如图 5—6 所示。

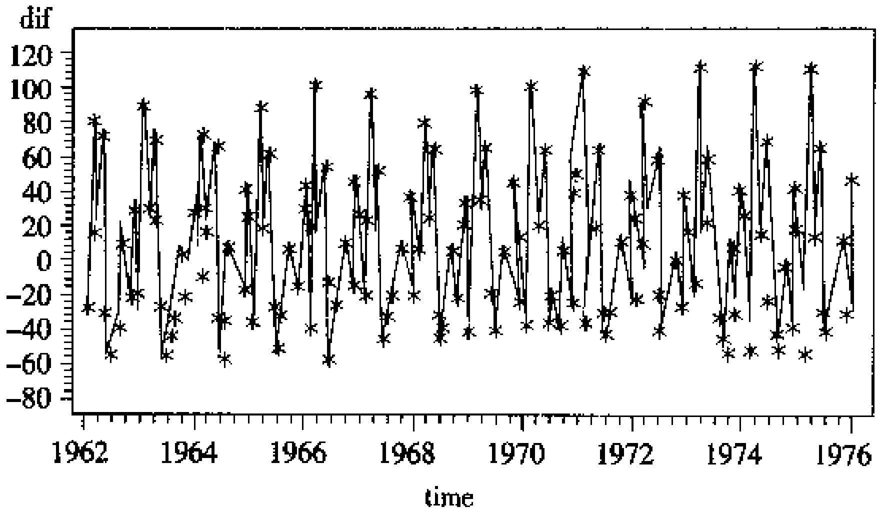


图 5—6　奶牛月产奶量 1 阶差分后序列时序图

图 5—6 显示，1 阶差分后线性递增信息被提取，1 阶差分后序列具有稳定的季节波动和随机波动。1 阶差分后序列再进行 12 步的周期差分，提取季节波动信息：

$$\nabla_{12}\nabla x_t = \nabla x_t - \nabla x_{t-12}$$

周期差分后序列 $\{\nabla_{12}\nabla x_t\}$ 时序图如图 5—7 所示。

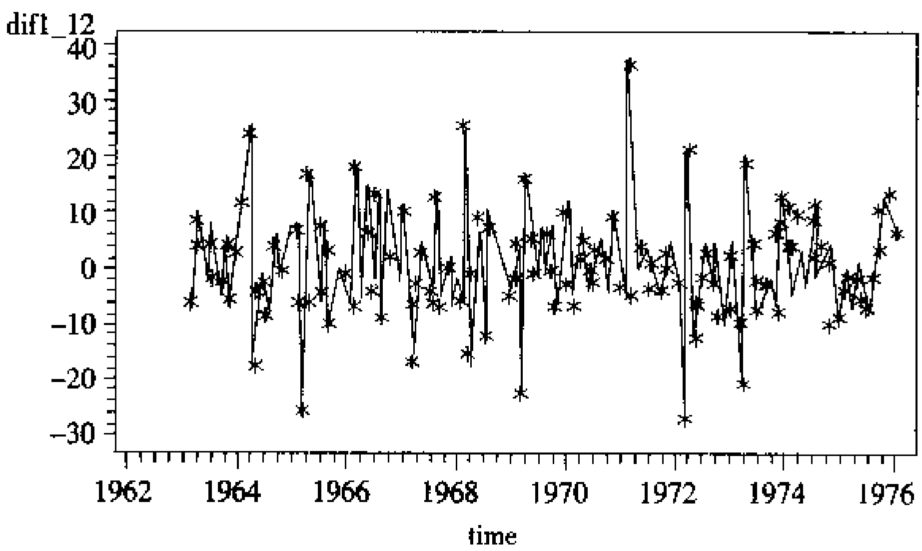


图 5—7　每头奶牛月产奶量 1 阶 12 步差分后序列时序图

图 5 -7 显示，周期差分可以非常好地提取周期信息。至此，差分运算已经比较充分地提取了原序列中蕴含的季节效应和长期趋势效应等确定性信息。差分后序列呈现典型的随机波动特征。

5.1.3 过差分

从理论上而言，足够多次的差分运算可以充分地提取原序列中的非平稳确定性信息。但应当注意的是，差分运算的阶数并不是越多越好。因为差分运算是一种对信息的提取、加工过程，每次差分都会有信息的损失，所以在实际应用中差分运算的阶数要适当，应当避免过度差分，简称过差分的现象。

【例 5.4】 假定线性非平稳序列 $\{x_t\}$ 形如

$$x_t=\beta_0+\beta_1 t+a_t$$

式中，$E(a_t)=0$，$\mathrm{Var}(a_t)=\sigma^2$，$\mathrm{Cov}(a_t, a_{t-1})=0$，$\forall i\geqslant 1$。

对 x_t 作 1 阶差分：

$$\begin{aligned}\nabla x_t &= x_t - x_{t-1}\\ &=\beta_0+\beta_1 t+a_t-[\beta_0+\beta_1(t-1)+a_{t-1}]\\ &=\beta_1+a_t-a_{t-1}\end{aligned}$$

显然差分后序列$\{\nabla x_t\}$为平稳序列。这说明 1 阶差分运算有效地提取了$\{x_t\}$中的非平稳确定性信息。

对 1 阶差分后序列∇x_t 再作一次差分：

$$\begin{aligned}\nabla^2 x_t &=\nabla x_t-\nabla x_{t-1}\\ &=\beta_1+a_t-a_{t-1}-[\beta_1+a_{t-1}-a_{t-2}]\\ &=a_t-2a_{t-1}+a_{t-2}\end{aligned}$$

显然 2 阶差分后序列$\{\nabla^2 x_t\}$也是平稳序列，它也将原序列中的非平稳趋势提取充分了。

考察它们的方差状况：

$$\mathrm{Var}(\nabla x_t)=\mathrm{Var}(a_t-a_{t-1})=2\sigma^2$$

$$\mathrm{Var}(\nabla^2 x_t)=\mathrm{Var}(a_t-2a_{t-1}+a_{t-2})=6\sigma^2$$

显然 2 阶差分后序列 $\{\nabla^2 x_t\}$ 的方差大于 1 阶差分后序列 $\{\nabla x_t\}$ 的方差，在这种场合下 2 阶差分就属于过差分。过差分实质上是因为过多次数的差分导致有效信息的无谓浪费而降低了估计的精度。

5.2 *ARIMA* 模型

差分运算具有强大的确定性信息提取能力，许多非平稳序列差分后会显示出平稳序列的性质，这时我们称这个非平稳序列为差分平稳序列。对差分平稳序列可以使用 *ARIMA* 模型进行拟合。

5.2.1 *ARIMA* 模型的结构

具有如下结构的模型称为求和自回归移动平均（autoregressive integrated moving average）模型，简记为 *ARIMA*（p，d，q）模型：

$$\begin{cases}\Phi(B)\nabla^d x_t=\Theta(B)\varepsilon_t \\ E(\varepsilon_t)=0,\mathrm{Var}(\varepsilon_t)=\sigma_\varepsilon^2,E(\varepsilon_t\varepsilon_s)=0,s\neq t \\ Ex_s\varepsilon_t=0,\forall s<t\end{cases} \tag{5.1}$$

式中：

$\nabla^d=(1-B)^d$

$\Phi(B)=1-\phi_1 B-\cdots-\phi_p B^p$，为平稳可逆 $ARMA(p,q)$模型的自回归系数多项式

$\Theta(B)=1-\theta_1 B-\cdots-\theta_q B^q$，为平稳可逆 $ARMA(p,q)$模型的移动平滑系数多项式

式(5.1)可以简记为：

$$\nabla^d x_t=\frac{\Theta(B)}{\Phi(B)}\varepsilon_t \tag{5.2}$$

式中，$\{\varepsilon_t\}$ 为零均值白噪声序列。

由式（5.2）显而易见，*ARIMA* 模型的实质就是差分运算与 *ARMA* 模型的组合。这一关系意义重大。这说明任何非平稳序列只要通过适当阶数的差分实现差分后平稳，就可以对差分后序列进行 *ARMA* 模型拟合了。而 *ARMA* 模型的分析方法非常成熟，这意味着对差分平稳序列的分析也将是非常简单、非常可靠的了。

求和自回归移动平均模型这个名字的由来是因为 d 阶差分后序列可以表示为：

$$\nabla^d x_t=\sum_{i=1}^{d}(-1)^d C_d^i x_{t-i}$$

式中，$C_d^i=\frac{d!}{i!(d-i)!}$，即差分后序列等于原序列的若干序列值的加权和，而对

它又可以拟合自回归移动平均（$ARMA$）模型，所以称它为求和自回归移动平均模型。

特别地，

当 $d=0$ 时，$ARIMA(p,d,q)$模型实际上就是 $ARMA(p,q)$模型。

当 $p=0$ 时，$ARIMA(0,d,q)$模型可以简记为 $IAM(d,q)$模型。

当 $q=0$ 时，$ARIMA(p,d,0)$模型可以简记为 $ARI(p,d)$模型。

当 $d=1$，$p=q=0$ 时，$ARIMA(0,1,0)$模型为：

$$\begin{cases} x_t = x_{t-1} + \varepsilon_t \\ E(\varepsilon_t)=0,\ \mathrm{Var}(\varepsilon_t)=\sigma_\varepsilon^2, E(\varepsilon_t\varepsilon_s)=0, s\neq t \\ Ex_s\varepsilon_t=0, \forall s<t \end{cases} \tag{5.3}$$

该模型被称为随机游走（random walk）模型，或醉汉模型。

随机游走模型的产生有一个有趣的典故。它最早于 1905 年 7 月由卡尔·皮尔逊（Karl Pearson）在《自然》杂志上作为一个问题提出：假如有个醉汉醉得非常严重，完全丧失方向感，把他放在荒郊野外，一段时间之后再去找他，在什么地方找到他的概率最大呢?

考虑到他完全丧失方向感，那么他第 t 步的位置将是他第 $t-1$ 步的位置再加一个完全随机的位移。用数学模型来描述任意时刻这个醉汉可能的位置，即为一个随机游走模型（5.3）。

1905 年 8 月，雷利爵士（Lord Rayleigh）对卡尔·皮尔逊的这个问题作出了解答。他算出这个醉汉离初始点的距离为 r 至 $r+\delta r$ 的概率为：

$$\frac{2}{nl^2}e^{-r^2/nl^2}r\delta r$$

且当 n 很大时，该醉汉离初始点的距离服从零均值正态分布。这意味着，假如有人想去寻找该醉汉的话，最好是去初始点附近找他，该地点是醉汉未来位置的无偏估计值。

作为一个最简单的 $ARIMA$ 模型，随机游走模型目前广泛应用于计量经济学领域。传统的经济学家普遍认为投机价格的走势类似于随机游走模型，随机游走模型也是有效市场理论（efficient market theory）的核心。

5.2.2 *ARIMA* 模型的性质

一、平稳性

假如$\{x_t\}$服从 $ARIMA(p, d, q)$模型：

$$\Phi(B)\nabla^d x_t = \Theta(B)\varepsilon_t$$

式中：

$$\nabla^d=(1-B)^d$$

$$\Phi(B)=1-\phi_1 B-\cdots-\phi_p B^p$$

$$\Theta(B)=1-\theta_1 B-\cdots-\theta_q B^q$$

记 $\varphi(B)=\Phi(B)\nabla^d$，$\varphi(B)$被称为广义自回归系数多项式。显然 $ARIMA$ 模型的平稳性完全由 $\varphi(B)=0$ 的根的性质决定。

因为 $\{x_t\}$ d 阶差分后平稳，服从 $ARMA$ (p, q) 模型，所以不妨设

$$\Phi(B)=\prod_{i-1}^{p}(1-\lambda_i B),\ |\lambda_i|<1;i=1,2,\cdots,p$$

则

$$\varphi(B)=\Phi(B)\nabla^d=\left[\prod_{i=1}^{p}(1-\lambda_i B)\right](1-B)^d \tag{5.4}$$

由式（5.4）容易判断，$ARIMA$ (p, d, q) 模型的广义自回归系数多项式共有 $p+d$个根，其中 p 个根$\frac{1}{\lambda_1}$，…，$\frac{1}{\lambda_p}$在单位圆外，d 个根在单位圆上。

自回归系数多项式的根即为特征根的倒数，所以 $ARIMA$ (p, d, q) 模型共有 $p+d$ 个特征根，其中 p 个在单位圆内，d 个在单位圆上。

因为有 d 个特征根在单位圆上而非单位圆内，所以当 $d\neq 0$ 时，$ARIMA$ (p, d, q) 模型不平稳。

【例 5.5】 拟合随机游走序列：$x_t=x_{t-1}+\varepsilon_t$，$\varepsilon_t \sim NID$ (0, 100)。时序图如图 5—8 所示。

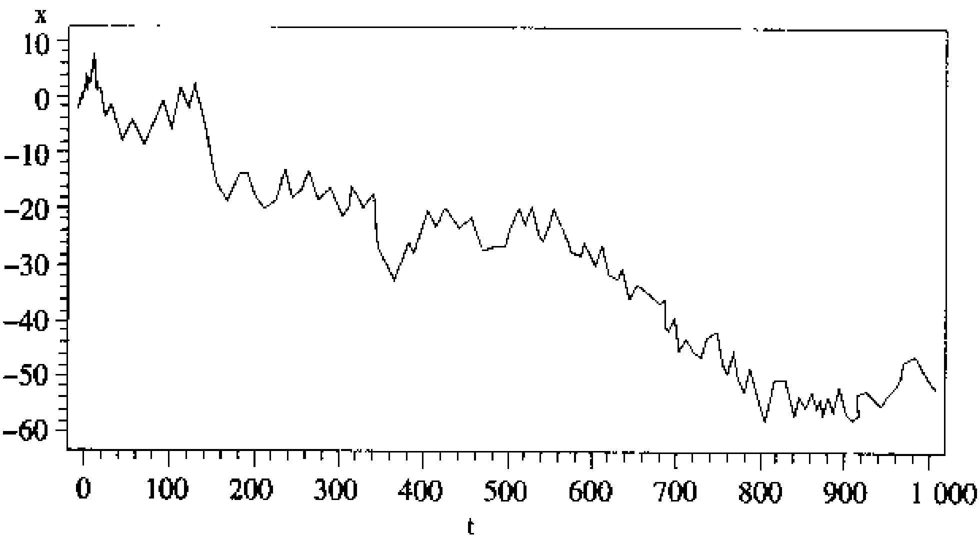


图 5—8 随机游走序列时序图

时序图清晰显示该序列均值非平稳。

二、方差齐性

对于 *ARIMA*（*p*，*d*，*q*）模型，当 $d\neq0$ 时，不仅均值非平稳，序列方差也非平稳。以最简单的随机游走模型 *ARIMA*（0，1，0）为例：

$$\begin{aligned} x_t &= x_{t-1}+\varepsilon_t \\ &= x_{t-2}+\varepsilon_t+\varepsilon_{t-1} \\ &\cdots\cdots \\ &= x_0+\varepsilon_t+\varepsilon_{t-1}+\cdots+\varepsilon_1 \end{aligned}$$

则

$$\mathrm{Var}(x_t)=\mathrm{Var}(x_0+\varepsilon_t+\varepsilon_{t-1}+\cdots+\varepsilon_1)=t\sigma_\varepsilon^2$$

这是一个时间 t 的递增函数，随着时间趋向无穷，序列$\{x_t\}$的方差也趋向无穷。

但 1 阶差分之后

$$\nabla x_t=\varepsilon_t$$

差分后序列方差齐性

$$\mathrm{Var}(\nabla x_t)=\sigma_\varepsilon^2$$

5.2.3 *ARIMA* 模型建模

在掌握了 *ARMA* 模型建模的方法之后，尝试使用 *ARIMA* 模型对观察序列建模是一件比较简单的事情。它遵循如下操作流程，如图 5—9 所示。

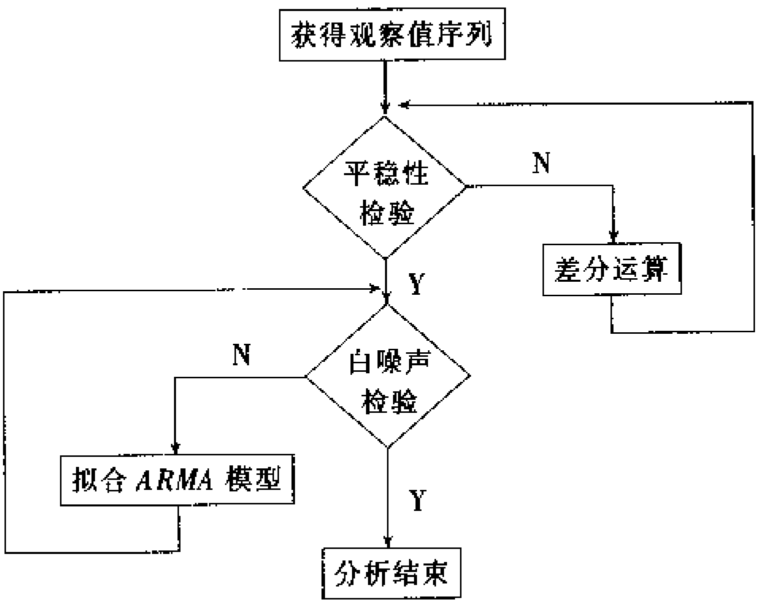


图 5—9 建模流程

根据这种建模流程，对一个真实序列建模。

【例 5.6】 对 1952—1988 年中国农业实际国民收入指数序列建模。

一、获得观察值序列

1952—1988 年中国农业实际国民收入指数见附录 1.14。

二、判断序列的平稳性

该序列时序图如图 5—10 所示。

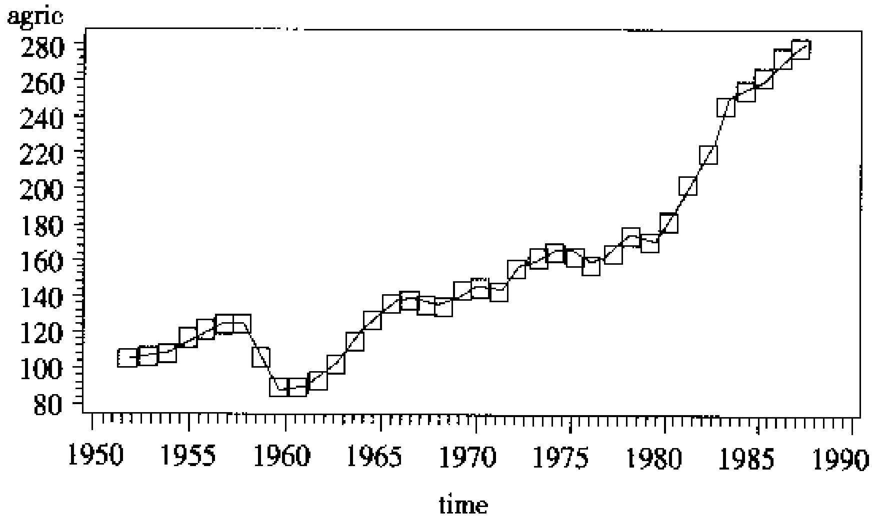


图 5—10 1952—1988 年中国农业实际国民收入指数时序图

时序图显示该序列有着显著的趋势，为典型的非平稳序列。

三、对原序列进行差分运算

因为原序列呈现出近似线性趋势，所以选择 1 阶差分。1 阶差分后序列时序图如图 5—11 所示。

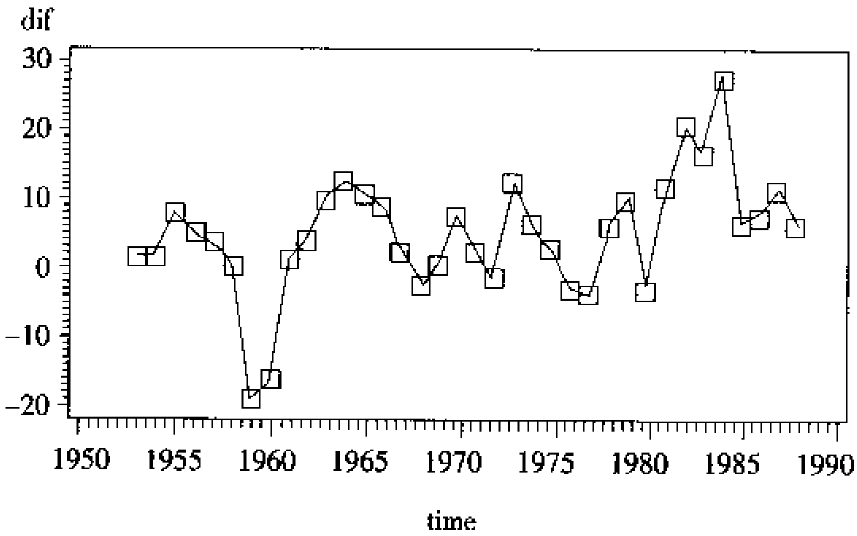


图 5—11 中图农业实际国民收入指数 1 阶差分后序列时序图

时序图显示出差分后序列在均值附近比较稳定地波动。为了进一步确定平稳性，考察差分后序列的自相关图，如图 5—12 所示。

```
                                  Autocorrelations
Lag  Covariance    Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
 0    78.239 167    1.000 00    |                    |********************|
 1    42.075 733    0.537 78    |             .      |***********         |
 2    16.246 605    0.207 65    |            .       |****   .            |
 3     7.058 588    0.090 22    |           .        |**      .           |
 4   -11.132 207    -.142 28    |           .     ***|        .           |
 5    -7.917 076    -.101 19    |           .      **|        .           |
 6    -9.245 185    -.118 17    |           .      **|        .           |
 7   -11.564 313    -.147 81    |           .     ***|        .           |
 8     7.108 735    0.090 86    |           .        |**      .           |
 9    12.965 116    0.165 71    |           .        |***     .           |
10    -0.909 105    -.011 62    |           .        |        .           |
11    -2.455 085    -.031 38    |           .       *|        .           |
12    -3.501 852    -.044 76    |           .       *|        .           |
13    -6.583 063    -.084 14    |           .      **|        .           |
14    -7.883 765    -.100 76    |           .      **|        .           |
15    -4.783 310    -.061 14    |           .       *|        .           |
16     2.087 515    0.026 68    |           .        |*       .           |
17    12.894 776    0.164 81    |           .        |***     .           |
18    15.631 250    0.199 79    |           .        |****    .           |

                          "·"marks two standard errors
```

图 5—12　中国农业实际国民收入指数 1 阶差分后序列自相关图

自相关图显示序列有很强的短期相关性，所以可以初步认为 1 阶差分后序列平稳。

四、对平稳的 1 阶差分序列进行白噪声检验

白噪声检验结果如表 5—1 所示。

表 5—1

1 阶差分后序列白噪声检验		
延迟阶数	χ^2 统计量	P 值
6	15.33	0.017 8
12	18.33	0.106 0
18	24.66	0.134 4

在检验的显著性水平取为 0.05 的条件下，由于延迟 6 阶的 χ^2 检验统计量的 P 值为0.017 8，小于 0.05，所以该差分后序列不能视为白噪声序列，即差分后序列还蕴含着不容忽视的相关信息可供提取。

五、对平稳非白噪声差分序列拟合 *ARMA* 模型

1 阶差分后序列的自相关图（图 5—12）已经显示出该序列有自相关系数 1 阶截尾的性质。再考察其偏自相关系数的性质。如图 5—13 所示。

```
                         Partial Autocorrelations
Lag    Correlation   -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  1     0.537 78      |                .    |***********         |
  2    -0.114 74      |                .  **|     .              |
  3     0.039 12      |                .    |*    .              |
  4    -0.270 03      |                .*****|    .              |
  5     0.162 19      |                .    |***  .              |
  6    -0.177 87      |                . ****|    .              |
  7     0.030 51      |                .    |*    .              |
  8     0.210 04      |                .    |**** .              |
  9     0.029 02      |                .    |*    .              |
 10    -0.257 95      |                .*****|    .              |
 11     0.044 21      |                .    |*    .              |
 12     0.043 46      |                .    |*    .              |
 13    -0.038 57      |                .   *|     .              |
 14    -0.155 91      |                .  ***|    .              |
 15     0.218 92      |                .    |**** .              |
 16     0.008 55      |                .    |     .              |
 17     0.054 96      |                .    |*    .              |
 18     0.018 25      |                .    |     .              |
```

图 5—13 中国农业实际国民收入指数 1 阶差分后序列偏自相关图

偏自相关图显示出显著的不截尾性，所以考虑用 *MA*（1）模型拟合 1 阶差分后序列。考虑到前面已经进行的 1 阶差分运算，实际上也就是用 *ARIMA*（0，1，1）模型拟合原序列。在条件最小二乘估计原理下，拟合结果为：

$$(1-B)x_t=4.996\ 61+(1+0.707\ 66B)\varepsilon_t,\mathrm{Var}(\varepsilon_t)=56.487\ 63$$

六、对残差序列进行检验

检验结果如表 5—2 所示。

表 5—2

残差白噪声检验			参数显著性检验		
延迟阶数	χ^2 统计量	*P* 值	待估参数	*t* 统计量	*P* 值
6	3.63	0.603 6	μ	2.39	0.022 3
12	7.86	0.726 2	θ_1	−5.58	<0.000 1
18	11.03	0.855 2			

显然，拟合检验统计量的 *P* 值都显著大于显著性检验水平 0.05，可以认为该残差序列即为白噪声序列，系数显著性检验显示两参数均显著。这说明 *ARIMA*（0，1，1)模型对该序列建模成功。

5.2.4 *ARIMA* 模型预测

在最小均方误差预测原理下，*ARIMA* 模型的预测和 *ARMA* 模型的预测方

法非常相似。

$ARIMA$（p，d，q）模型的一般表示方法为：

$$\Phi(B)(1-B)^d x_t=\Theta(B)\varepsilon_t$$

和 $ARMA$ 模型一样，也可以用历史观察值的线性函数表示它：

$$\begin{aligned}x_t &=\varepsilon_t+\Psi_1\varepsilon_{t-1}+\Psi_2\varepsilon_{t-2}+\cdots\\ &=\Psi(B)\varepsilon_t\end{aligned}$$

式中，Ψ_1，Ψ_2，…的值由如下等式确定：

$$\Phi(B)(1-B)^d\Psi(B)=\Theta(B)$$

如果把 $\Phi^*(B)$ 记为广义自相关函数，有

$$\Phi^*(B)=\Phi(B)(1-B)^d=1-\phi_1 B-\phi_2 B^2+\cdots$$

容易验证 Ψ_1，Ψ_2，…的值满足如下递推公式：

$$\begin{cases}\Psi_1=\phi_1-\theta_1\\ \Psi_2=\phi_1\Psi_1+\phi_2-\theta_2\\ \cdots\cdots\cdots\cdots\\ \Psi_j=\phi_1\Psi_{j-1}+\cdots+\phi_{p+d}\Psi_{j-p-d}-\theta_j\end{cases}$$

式中，$\Psi_j=\begin{cases}0, & j<0\\ 1, & j=0\end{cases}$，$\theta_j=0$，$j>q$

那么，x_{t+l}的真实值为：

$$x_{t+1}=(\varepsilon_{t+l}+\Psi_1\varepsilon_{t+l-1}+\cdots+\Psi_{l-1}\varepsilon_{t+1})+(\Psi_l\varepsilon_t+\Psi_{l+1}\varepsilon_{t-1}+\cdots)$$

由于 ε_{t+l}，ε_{t-l-1}，…，ε_{t+l}的不可获得性，所以 x_{t+1}的估计值只能为：

$$\hat{x}_t(l)=\Psi_0^*\varepsilon_t+\Psi_1^*\varepsilon_{t-1}+\Psi_2^*\varepsilon_{t-2}+\cdots$$

真实值与预报值之间的均方误差为

$$E[x_{t+l}-\hat{x}_t(l)]^2=(1+\Psi_1^2+\cdots+\Psi_{l-1}^2)\sigma_\varepsilon^2+\sum_{j=0}^{\infty}(\Psi_{l+j}-\Psi_j^*)^2\sigma_\varepsilon^2$$

要使均方误差最小，当且仅当

$$\Psi_j^*=\Psi_{l+j}$$

所以在均方误差最小原则下，l 期预报值为：

$$\hat{x}_t=(l)=\Psi_l\varepsilon_t+\Psi_{l+1}\varepsilon_{t-1}+\Psi_{l+2}\varepsilon_{t-2}+\cdots$$

l 期预报误差为：

$$e_t(l)=\varepsilon_{t+l}+\Psi_1\varepsilon_{t+l-1}+\cdots+\Psi_{l-1}\varepsilon_{t+1}$$

真实值等于预报值加上预报误差：

$$\begin{aligned}x_{t+l}&=(\varepsilon_{t+l}+\Psi_1\varepsilon_{t+l-1}+\cdots+\Psi_{l-1}\varepsilon_{t+1})+(\Psi_l\varepsilon_t+\Psi_{l+1}\varepsilon_{t-1}+\cdots)\\ &=e_t(l)+\hat{x}_t(l)\end{aligned}$$

l 期预报误差的方差为：

$$\mathrm{Var}[e_t(l)]=(1+\Psi_1^2+\cdots+\Psi_{l-1}^2)\sigma_\varepsilon^2$$

【例 5.7】 已知 *ARIMA* (1, 1, 1) 模型为$(1-0.8B)(1-B)x_t=(1-0.6B)\varepsilon_t$，且 $x_{t-1}=4.5$，$x_t=5.3$，$\varepsilon_t=0.8$，$\sigma_\varepsilon^2=1$。求 x_{t+3}的 95%的置信区间。

展开原模型，等价形式为：

$$(1-1.8B+0.8B^2)x_t=(1-0.6B)\varepsilon_t$$

$$x_t=1.8x_{t-1}-0.8x_{t-2}+\varepsilon_t-0.6\varepsilon_{t-1}$$

则预报值的递推公式为：

$$\hat{x}_t(1)=1.8x_t-0.8x_{t-1}-0.6\varepsilon_t=5.46$$

$$\hat{x}_t(2)=1.8\hat{x}_t(1)-0.8x_t=5.59$$

$$\hat{x}_t(3)=1.8\hat{x}_t(2)-0.8\hat{x}_t(1)=5.69$$

3 期预报误差的方差为：

$$\mathrm{Var}[e(3)]=(1+\Psi_1^2+\Psi_2^2)\sigma_\varepsilon^2$$

根据 $\Phi(B)(1-B)^d\Psi(B)=\Theta(B)$，有

$$(1-0.8B)(1-B)(1+\Psi_1B+\Psi_2B^2+\cdots)=1-0.6B_t$$

根据 B 的同幂次系数相等的原理，可以确定 Ψ_1，Ψ_2 的值：

$$\begin{cases}\Psi_1=1.8-0.6=1.2\\ \Psi_2=1.8\Psi_1-0.8=1.36\end{cases}$$

直接用 Ψ_1，Ψ_2 的递推公式可以更快地得到相同的结果，则

$$\mathrm{Var}[e(3)]=(1+\Psi_1^2+\Psi_2^2)\sigma_\varepsilon^2=4.2896$$

x_{t+3}的 95%的置信区间为$(\hat{x}_t(3)-1.96\sqrt{\mathrm{Var}(e(3))},\ \hat{x}_t(3)+1.96\sqrt{\mathrm{Var}(e(3))})$即(1.63,9.75)。

【例 5.6 续】 对 1952—1988 年中国农业实际国民收入指数序列做为期 10 年的预测。如农 5—3 所示。

表 5—3

年份	预测值	标准差	95%置信下限	95%置信上限
1989	285.517 8	7.515 8	270.787 1	300.248 6
1990	290.514 5	14.873 2	261.363 6	319.665 3
1991	295.511 1	19.645 2	257.007 1	334.015 0
1992	300.507 7	23.466 1	254.514 9	346.500 4
1993	305.504 3	26.746 6	253.081 9	357.926 7
1994	310.500 9	29.666 6	252.355 5	368.646 3
1995	315.497 5	32.323 8	252.144 0	378.851 0
1996	320.494 1	34.778 6	252.329 3	388.659 0
1997	325.490 7	37.071 2	252.832 4	398.149 0
1998	330.487 3	39.230 1	253.597 8	407.376 9

预测图如图 5—14 所示。

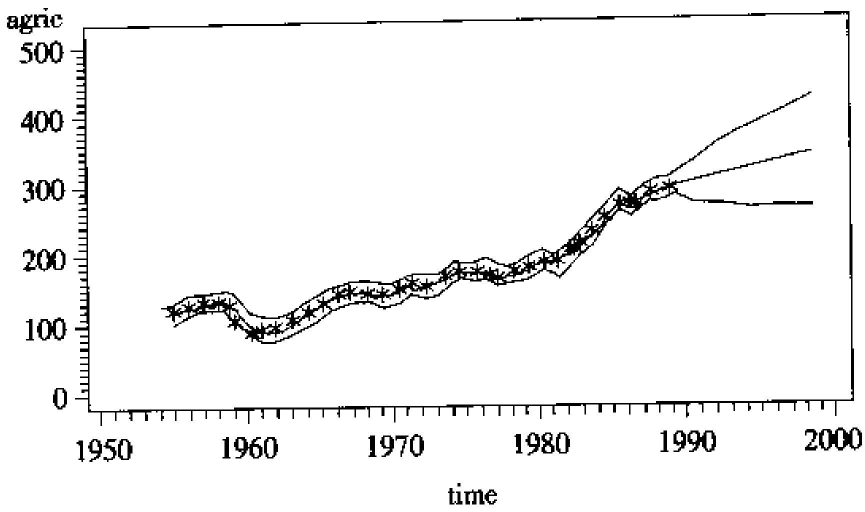


图 5—14　中国农业实际国民收入指数序列预测图

图中，星号为序列观察值；中间曲线为序列预测值；上下曲线为 95%的置信区间。随着预测时期变长，预测误差越来越大，预测区间呈现为喇叭型，这是时间序列预测的典型特点。

5.2.5　疏系数模型

ARIMA（p，d，q）模型是指 d 阶差分后自相关最高阶数为 p，移动平均最高阶数为 q 的模型，通常它包含 $p+q$ 个独立的未知系数：ϕ_1，…，ϕ_p，θ_1，…，θ_q。

如果该模型中有部分自相关系数 ϕ_j（$1\leqslant j<p$）或部分移动平滑系数 θ_k（$1\leqslant k<q$）为零，即原 *ARIMA*（p，d，q）模型中有部分系数省缺了，那么该模型称为疏系数模型。

如果只是自相关部分有省缺系数，那么该疏系数模型可以简记为：

$$ARIMA((p_1,\cdots,p_m),d,q)$$

式中，p_1，…，p_m 为非零自相关系数的阶数。

如果只是移动平滑部分有省缺系数，那么该疏系数模型可以简记为：

$$ARIMA(p,d,(q_1,\cdots,q_n))$$

式中，q_1，…，q_n 为非零移动平滑系数的阶数。

如果自相关和移动平滑部分都有省缺，可以简记为

$$ARIMA((p_1,\cdots,p_m),d,(q_1,\cdots,q_n))$$

在实际操作中，疏系数模型时有应用。

【例 5.8】 对 1917—1975 年美国 23 岁妇女每万人生育率序列建模（数据见附录 1.15）。

一、绘制时序图

如图 5—15 所示。

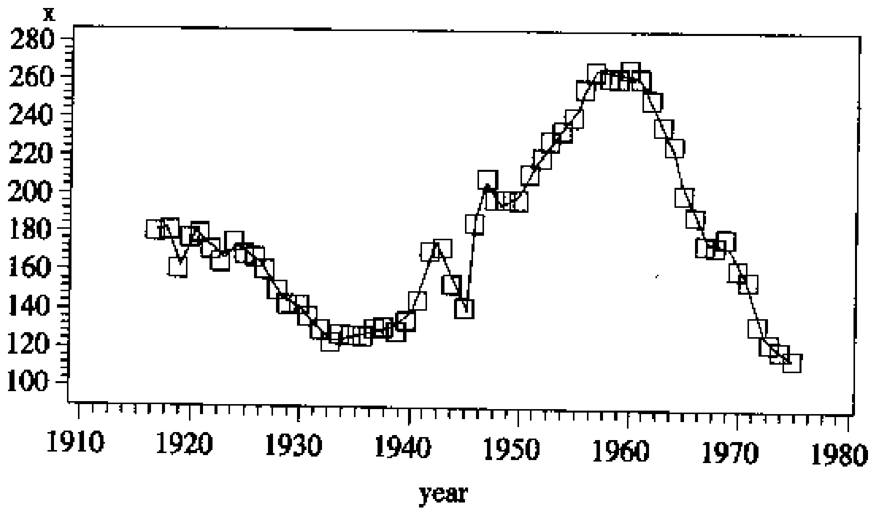


图 5—15　美国 23 岁妇女每万人生育率序列时序图

二、差分平稳

时序图显示，序列具有长期趋势，对序列进行 1 阶差分 $\nabla x_t = x_t - x_{t-1}$，观察差分后序列的 ∇x_t 时序图。如图 5—16 所示。

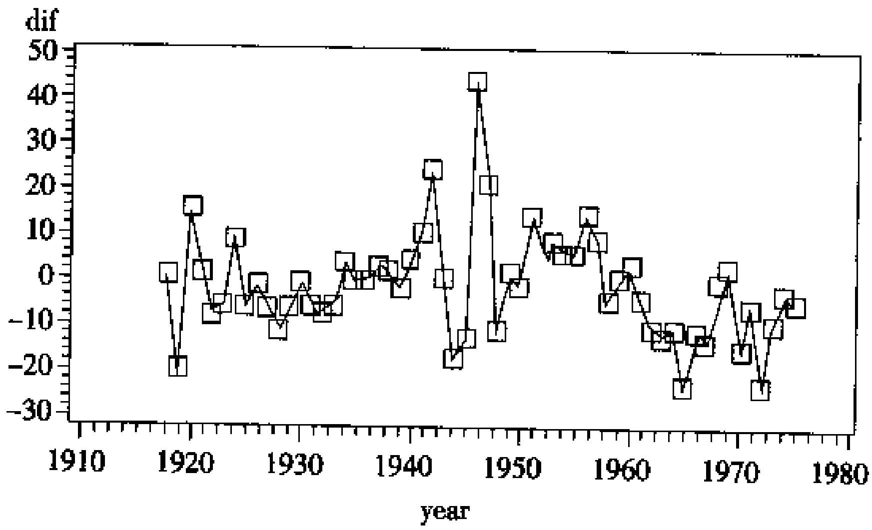


图 5—16　美国 23 岁妇女每万人生育率 1 阶差分后序列时序图

时序图显示长期趋势信息基本被差分运算提取充分，考察差分后序列的自相关图，进一步确定平稳性。如图 5—17 所示。

```
                                   Autocorrelations
Lag  Covariance   Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
 0     145.928     1.000 00     |                    |********************|
 1   41.471 097    0.284 19     |                .   |******              |
 2   -6.153 942    -.042 17     |               .   *|     .              |
 3   11.679 824    0.080 04     |               .    |**   .              |
 4   47.775 344    0.327 39     |               .    |*******             |
 5   47.552 927    0.325 86     |               .    |*******             |
 6   21.691 431    0.148 64     |              .     |***    .            |
 7    9.102 516    0.062 38     |              .     |*      .            |
 8    3.544 593    0.024 29     |              .     |       .            |
 9   12.658 538    0.086 74     |              .     |**     .            |
10   13.544 854    0.092 82     |              .     |**     .            |
11   -5.308 811    -.036 38     |              .    *|       .            |
12  -10.154 165    -.069 58     |              .    *|       .            |
13   -6.521 036    -.044 69     |              .    *|       .            |
14   -7.235 034    -.049 58     |              .    *|       .            |
15  -16.998 688    -.116 49     |              .   **|       .            |
16  -18.817 139    -.128 95     |              .  ***|       .            |
17  -10.170 169    -.069 69     |              .    *|       .            |
18  -21.791 201    -.149 33     |              .  ***|       .            |

                          "·"marks two standard errors
```

图 5—17　美国 23 岁妇女每万人生育率 1 阶差分后序列自相关图

自相关图显示延迟 6 阶之后，自相关系数基本都在零值附近波动。可以认为自相关系数具有短期相关性。该差分后序列平稳。

三、模型定阶

为了确定模型的阶数，我们还需考察偏自相关图，如图 5—18 所示。

```
                       Partial Autocorrelations
Lag   Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
 1     0.284 19     |               .    |******              |
 2    -0.133 73     |               . ***|    .               |
 3     0.145 80     |               .    |*** .               |
 4     0.284 05     |               .    |******              |
 5     0.196 04     |               .    |****.               |
 6     0.079 73     |               .    |**  .               |
 7     0.029 65     |               .    |*   .               |
 8    -0.106 04     |               .  **|    .               |
 9    -0.050 98     |               .   *|    .               |
10    -0.060 54     |               .   *|    .               |
11    -0.146 40     |               . ***|    .               |
12    -0.059 23     |               .   *|    .               |
13    -0.050 89     |               .   *|    .               |
14    -0.062 81     |               .   *|    .               |
15    -0.074 25     |               .   *|    .               |
16    -0.033 83     |               .   *|    .               |
17     0.027 52     |               .    |*   .               |
18    -0.089 11     |               .  **|    .               |
```

图 5—18　美国 23 岁妇女每万人生育率 1 阶差分后序列偏自相关国

偏自相关图显示，除了延迟 1 阶和 4 阶的偏自相关系数显著大于 2 倍标准差间之外，其他阶数的偏自相关系数都比较小。

根据自相关图和偏自相关图的这个特点，进行模型的定阶。考虑到偏自相关图中只有延迟 1 阶和 4 阶的偏自相关系数显著大于 2 倍标准差，所以考虑构造疏系数模型 $AR(1,4)$。

综合考虑前面的差分运算，实际上是对原序列拟合疏系数模型 $ARIMA$ ((1，4)，1，0)。

四、参数估计

使用条件最小二乘估计，确定模型的口径为：

$$(1-B)x_t=\frac{1}{1-0.266\,33B-0.335\,97B^4}\varepsilon_t$$

五、模型检验

模型检验如表 5—4 所示。

表 5—4

残差白噪声检验			参数显著性检验		
延迟阶数	χ^2 统计量	P 值	待估参数	t 统计量	P 值
6	4.26	0.371 8	ϕ_1	2.2	0.031 8
12	5.72	0.838 4	ϕ_4	2.67	0.010 0

残差检验结果显示残差序列可视为白噪声序列。参数显著性检验结果显示两参数均高度显著，所以该疏系数模型拟合成功。

在进行模型定阶时，如果没有经验不敢直接构造疏系数模型，我们也可以通过传统的定阶方法，通过反复尝试和删减不显著参数得到相同的疏系数模型。

比如本例中，在模型定阶阶段可能会有如下考虑和选择，如表 5—5 所示。

表 5—5

考虑	选择模型	拟合结果
自相关系数 6 阶截尾	$MA(6)$	残差不能通过白噪声检验 参数 θ_1，θ_2，θ_2，θ_6 均不显著
偏自相关系数 4 阶截尾	$AR(4)$	残差通过白噪声检验 参数 ϕ_2，ϕ_3 不显著
自相关和偏自相关截尾的阶数都偏长	$ARMA(1，1)$	残差不能通过白噪声检验 两参数 θ_1，ϕ_1 均不显著

因为只有 AR 模型的残差通过了白噪声检验，只是拟合的参数过多，有部分参数不显著。删除不显著的参数 ϕ_2，ϕ_3，优化模型。通过这一系列的操作，最后殊途同归，得到的是同一个疏系数模型。

5.2.6 季节模型

ARIMA 模型可以对具有季节效应的序列建模。根据季节效应提取的难易程度，可以分为简单季节模型和乘积季节模型。

一、简单季节模型

简单季节模型是指序列中的季节效应和其他效应之间是加法关系，即

$$x_t = S_t + T_t + I_t$$

这时，各种效应信息的提取都非常容易。通常简单的周期步长差分即可将序列中的季节信息提取充分，简单的低阶差分即可将趋势信息提取充分，提取完季节信息和趋势信息之后的残差序列就是一个平稳序列，可以用 *ARMA* 模型拟合。

所以简单季节模型实际上就是趋势差分、季节差分之后序列即可转化为平稳，再对差分后的平稳序列进行拟合。它的模型结构通常如下：

$$\nabla_D \nabla^d x_t = \frac{\Theta(B)}{\Phi(B)} \varepsilon_t$$

式中：

(1) D 为周期步长，d 为提取趋势信息所用的差分阶数。

(2) $\{\varepsilon_t\}$ 为白噪声序列，且 $E(\varepsilon_t)=0$，$\mathrm{Var}(\varepsilon_t)=\sigma_\varepsilon^2$。

(3) $\Theta(B)=1-\theta_1 B-\theta_q B^q$，为 q 阶移动平均系数多项式。

(4) $\Phi(B)=1-\phi_1 B-\phi_p B^p$，为 p 阶自回归系数多项式。

【例 5.9】 拟合 1962—1991 年德国工人季度失业率序列（数据见附录 1.16）。

1. 绘制观察值序列时序图

时序图如图 5—19 所示。

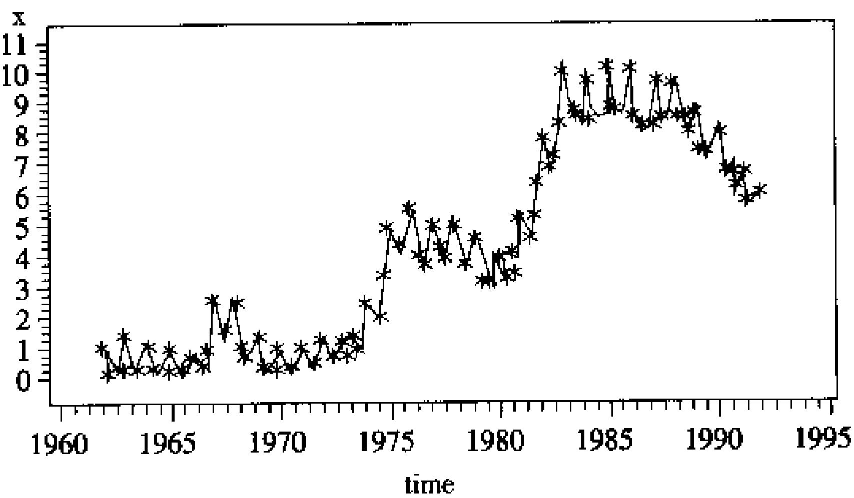


图 5—19 德国工人季度失业率序列时序图

时序图显示该序列既含有长期趋势又含有以年为周期的季节效应。

2. 差分平稳化

对原序列作1阶差分消除趋势，再作4步差分消除季节效应的影响，差分后序列的时序图如图5—20所示。

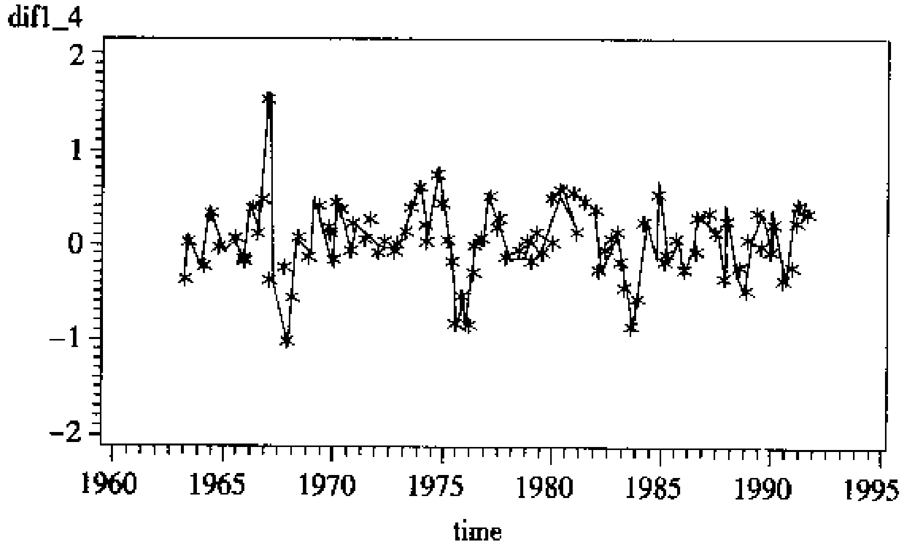


图5—20 德国工人季度失业率1阶4步差分后序列时序图

时序图显示差分后序列已无显著趋势或周期，随机波动比较平稳。

考察差分后序列的随机性，白噪声检验结果如表5—6所示。

表5—6

白噪声检验		
延迟阶数	χ^2统计量	P值
6	43.84	<0.000 1
12	51.71	<0.000 1
18	54.48	<0.000 1

检验结果显示，差分后序列蕴含着很强的相关信息，不能视为白噪声序列。需要对差分后序列进一步拟合*ARMA*模型。

3. 模型拟合

考察差分后序列的自相关图和偏自相关图的性质，为拟合模型定阶。如图5—21和图5—22所示。

自相关图显示差分后序列仍含有一定的季节效应，所以延迟4阶之后，自相关系数又有一个反弹。由于延迟1阶到3阶，延迟4阶到7阶衰减得非常迅速，所以该序列具有短期相关性。

```
                                  Autocorrelations

Lag  Covariance    Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
 0    0.124 721    1.000 00    |                    |********************|
 1    0.052 641    0.422 07    |                .   |********            |
 2    0.023 569    0.188 98    |                .   |****                |
 3    0.009 040 4  0.072 48    |                .   |*  .                |
 4   -0.029 789    -.238 84    |               *****|   .                |
 5   -0.029 783    -.238 79    |               *****|    .               |
 6   -0.022 559    -.180 87    |               .****|    .               |
 7   -0.024 312    -.194 93    |               .****|    .               |
 8   -0.011 082    -.088 85    |               .  **|    .               |
 9   -0.007 450 5  -.059 74    |               .   *|    .               |
10    0.002 716 0  0.021 78    |               .    |    .               |
11    0.002 715 5  0.021 77    |               .    |    .               |
12   -0.013 672    -.109 62    |               .  **|    .               |
13   -0.002 542 3  -.020 38    |               .    |    .               |
14   -0.003 398 8  -.027 25    |               .   *|    .               |
15   -0.015 005    -.120 31    |               .  **|    .               |
16    0.002 136 7  0.017 13    |               .    |    .               |
17    0.007 026 1  0.056 33    |               .    |*   .               |
18   -0.004 619 8  -.037 04    |               .   *|    .               |

                          "."marks two standard errors
```

图 5—21 德国工人季度失业率 1 阶 4 步差分后序列自相关图

```
                          Partial Autocorrelations

Lag   Correlation   -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
 1     0.422 07    |                .   |********            |
 2     0.013 19    |                .   |   .                |
 3    -0.014 35    |                .   |   .                |
 4    -0.324 52    |              ******|   .                |
 5    -0.038 22    |                .  *|   .                |
 6    -0.017 21    |                .   |   .                |
 7    -0.070 52    |                .  *|   .                |
 8    -0.033 61    |                .  *|   .                |
 9    -0.079 38    |                . **|   .                |
10     0.056 48    |                .   |*  .                |
11    -0.077 72    |                . **|   .                |
12    -0.193 44    |                ****|   .                |
13     0.041 33    |                .   |*  .                |
14    -0.018 02    |                .   |   .                |
15    -0.127 41    |                .***|   .                |
16     0.012 61    |                .   |   .                |
17     0.037 10    |                .   |*  .                |
18    -0.115 50    |                . **|   .                |
```

图 5—22 德国工人季度失业率 1 阶 4 步差分后序列偏自相关图

偏自相关图显示，除了延迟 1 阶和 4 阶的偏自相关系数显著大于 2 倍标准差之外，其他阶数的偏自相关系数基本都在 2 倍标准差范围内波动。所以尝试拟合

疏系数模型 $AR(1, 4)$。使用条件最小二乘估计法，得到未知参数的估计值为：

$$\hat{\phi}_1=0.44746, \quad \hat{\phi}_4=-0.28132$$

综合考虑前面的差分运算，对德国工人季度失业率序列拟合的模型为：

$$(1-B)(1-B^4)x_t=\frac{1}{1-0.44746B+0.28132B^4}\varepsilon_t$$

4. 模型检验

对该模型进行检验，结果如表 5—7 所示。

表 5—7

残差白噪声检验			参数显著性检验		
延迟阶数	χ^2 统计量	P 值	待估参数	t 统计量	P 值
6	2.09	0.719 1	θ_1	5.48	<0.000 1
12	10.99	0.358 4	θ_4	−3.41	0.000 9

检验结果显示，残差序列通过白噪声检验，参数显著性检验显示两个参数均显著。这说明该模型拟合良好，对序列相关信息的提取充分。

将序列拟合值和序列观察值联合作图，通过图示也可以直观地看出该模型对序列的拟合效果良好。如图 5—23 所示。

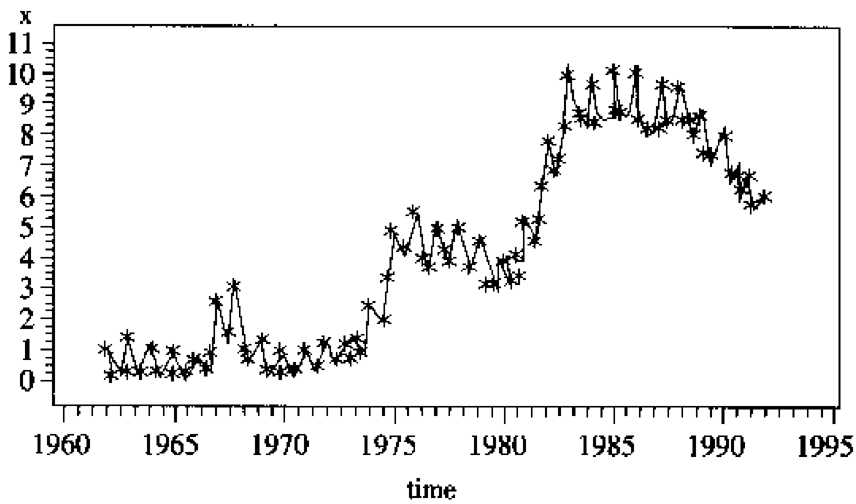


图 5—23　德国工人季度失业率序列模型拟合效果图

图中，星号为序列观察值；曲线为模型拟合值。

二、乘积季节模型

例 5.9 是一个既含有季节效应又含有长期趋势效应的简单序列，说它简单是因为这种序列季节效应、趋势效应和随机波动彼此之间很容易分开，这时简单的 *ARIMA* 模型即可拟合该序列的发展。

但更为常见的情况是，序列的季节效应、长期趋势效应和随机波动之间有着

复杂的相互纠缠关系，简单的 *ARIMA* 模型并不足以提取其中的相关关系，这时通常需要采用乘积季节模型。

【例 5.10】 1948—1981 年美国女性（大于 20 岁）月度失业率序列（数据见附录 1.17）。

1. 序列时序图

时序图如图 5—24 所示。

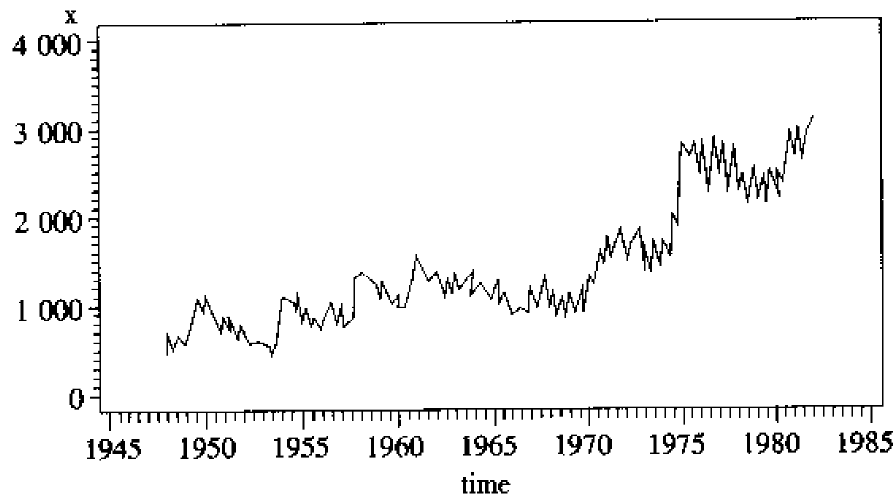


图 5—24 美国女性月度失业率序列时序图

时序图显示该序列具有长期递增趋势和以年为周期的季节效应。

2. 差分平稳

对原序列作 1 阶、12 步差分，希望提取原序列趋势效应和季节效应，差分后序列时序图如图 5—25 所示。

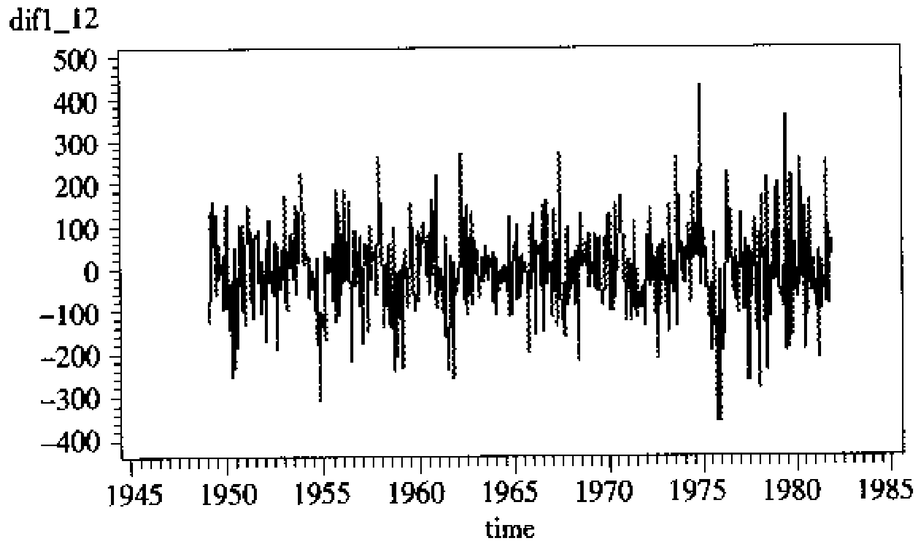


图 5—25 美国女性月度失业率 1 阶 12 步差分后序列时序图

时序图显示差分后序列类似平稳。

3. 模型定阶

考察差分后序列自相关图的性质，进一步确定平稳性判断，并估计拟合模型的阶数。如图 5—26 所示。

```
                                        Autocorrelations

Lag   Covariance   Correlation   -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1

  0   12 557.531    1.000 00     |                    |********************|
  1   -1 734.813    -.138 15     |                 ***| .                  |
  2    2 375.783    0.189 19     |                  . |****                |
  3      279.359    0.022 25     |                  . | .                  |
  4      759.808    0.060 51     |                  . |*.                  |
  5      138.150    0.011 00     |                  . | .                  |
  6      655.488    0.052 20     |                  . |*.                  |
  7   -1 028.202    -.081 88     |                  **| .                  |
  8      533.734    0.042 50     |                  . |*.                  |
  9     -158.605    -.012 63     |                  . | .                  |
 10   -1 606.237    -.127 91     |                 ***| .                  |
 11      858.206    0.068 34     |                  . |*.                  |
 12   -5 698.457    -.453 79     |           *********| .                  |
 13      521.120    0.041 50     |                 .  |* .                 |
 14     -509.219    -.040 55     |                 . *|  .                 |
 15   -1 020.660    -.081 28     |                 .**|  .                 |
 16     -730.212    -.058 15     |                 . *|  .                 |
 17      429.071    0.034 17     |                 .  |* .                 |
 18     -825.235    -.065 72     |                 . *|  .                 |
 19      592.947    0.047 22     |                 .  |* .                 |
 20     -565.282    -.045 02     |                 . *|  .                 |
 21      206.681    0.016 46     |                 .  |  .                 |
 22     -117.966    -.009 39     |                 .  |  .                 |
 23      774.691    0.061 69     |                 .  |* .                 |
 24     -929.421    -.074 01     |                 . *|  .                 |

                               "."marks two standard errors
```

图 5—26 美国女性月度失业率 1 阶 12 步差分后序列自相关图

自相关图显示延迟 12 阶自相关系数显著大于 2 倍标准差范围，这说明差分后序列中仍蕴含着非常显著的季节效应。延迟 1 阶、2 阶的自相关系数也大于 2 倍标准差，这说明差分后序列还具有短期相关性。

观察偏自相关图得到的结论和上面的结论一致。如图 5—27 所示。

根据差分后序列的自相关图和偏自相关图的性质，尝试拟合如下 *ARMA* 模型，拟合效果均不理想，拟合残差均通不过白噪声检验。如表 5—8 所示。

```
                         Partial Autocorrelations
Lag  Correlation   -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  1   -0.138 15    |                 ***| .                  |
  2    0.173 42    |                  . |***                 |
  3    0.071 27    |                  . |*.                  |
  4    0.041 01    |                  . |*.                  |
  5    0.006 52    |                  . | .                  |
  6    0.036 43    |                  . |*.                  |
  7   -0.083 99    |                  **| .                  |
  8    0.002 91    |                  . | .                  |
  9    0.017 76    |                  . | .                  |
 10   -0.140 08    |                 ***| .                  |
 11    0.041 84    |                  . |*.                  |
 12   -0.428 81    |           *********| .                  |
 13   -0.073 13    |                  .*| .                  |
 14    0.118 56    |                  . |**                  |
 15   -0.041 73    |                  .*| .                  |
 16   -0.032 32    |                  .*| .                  |
 17    0.043 20    |                  . |*.                  |
 18    0.007 66    |                  . | .                  |
 19   -0.032 05    |                  .*| .                  |
 20   -0.008 93    |                  . | .                  |
 21    0.035 29    |                  . |*.                  |
 22   -0.121 75    |                  **| .                  |
 23    0.093 76    |                  . |**                  |
 24   -0.300 02    |              ******| .                  |
```

图 5—27　美国女性月度失业率 1 阶 12 步差分后序列偏自相关图

表 5—8

延迟阶数	拟合模型残差白噪声检验					
	AR(1,12)		*MA*(1,2,12)		*ARMA*(1,12)(1,12)	
	χ^2 值	*P* 值	χ^2 值	*P* 值	χ^2 值	*P* 值
6	14.58	0.005 7	9.5	0.023 3	15.77	0.000 4
12	16.42	0.088 3	14.19	0.115 8	17.99	0.021 3

反复地尝试结果均不理想，说明简单的 *ARMA* 模型并不适合于拟合这个序列。考虑到该序列既具有短期相关性又具有季节效应，短期相关性和季节效应不能简单地、可加性地提取，因而估计该序列的季节效应和短期相关性之间具有复杂的关联性。这时，通常假定短期相关性和季节效应之间具有乘积关系，尝试使用乘积模型来拟合序列的发展。

乘积模型的构造原理如下：

当序列具有短期相关性时，通常可以使用低阶 $ARMA(p, q)$ 模型提取。

当序列具有季节效应，季节效应本身还具有相关性时，季节相关性可以使用以周期步长为单位的 $ARMA(P, Q)$ 模型提取。

由于短期相关性和季节效应之间具有乘积关系，所以拟合模型实质为 $ARMA(p, q)$和 $ARMA(P, Q)$ 的乘积。综合前面的 d 阶趋势差分和 D 阶以周期 S 为步长的季节差分运算，对原观察值序列拟合的乘积模型完整的结构如下：

$$\nabla^d \nabla_S^D x_t = \frac{\Theta(B)\Theta_S(B)}{\Phi(B)\Phi_S(B)}\varepsilon_t$$

式中：

$$\Theta(B) = 1 - \theta_1 B - \cdots - \phi_q B^q$$

$$\Phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$$

$$\Theta_S(B) = 1 - \theta_1 B^S - \cdots - \phi_Q B^{QS}$$

$$\Phi_S(B) = 1 - \phi_1 B^S - \cdots - \phi_P B^{PS}$$

该乘积模型简记为 $ARIMA(p,d,q)\times(P,D,Q)_S$。

回到例 5.10 的模型定阶阶段，考虑到差分后序列短期相关性显著，尝试拟合乘积模型 $ARIMA(1,1,1)\times(0,1,1)_{12}$：

$$\nabla\nabla_{12} x_t = \frac{1-\theta_1 B}{1-\phi_1 B}(1-\theta_{12}B^{12})\varepsilon_t$$

4. 参数估计

使用条件最小二乘估计方法，确定该拟合模型的口径为：

$$\nabla\nabla_{12} x_t = \frac{1+0.661\ 37B}{1+0.789\ 78B}(1-0.773\ 94B^{12})\varepsilon_t$$

5. 模型检验

对拟合模型进行检验，检验结果显示该模型顺利通过残差白噪声检验和参数显著性检验。如表 5—9 所示。

表 5—9

残差白噪声检验			参数显著性检验		
延迟阶数	χ^2 统计量	P 值	待估参数	t 统计量	P 值
6	4.50	0.212 0	θ_1	−4.66	<0.000 1
12	9.41	0.400 2	θ_{12}	23.03	<0.000 1
18	20.58	0.150 7	ϕ_1	−6.81	<0.000 1

将序列拟合值和序列观察值联合作图，可以直观地看出该乘积模型对原序列的拟合效果良好。如图 5—28 所示。

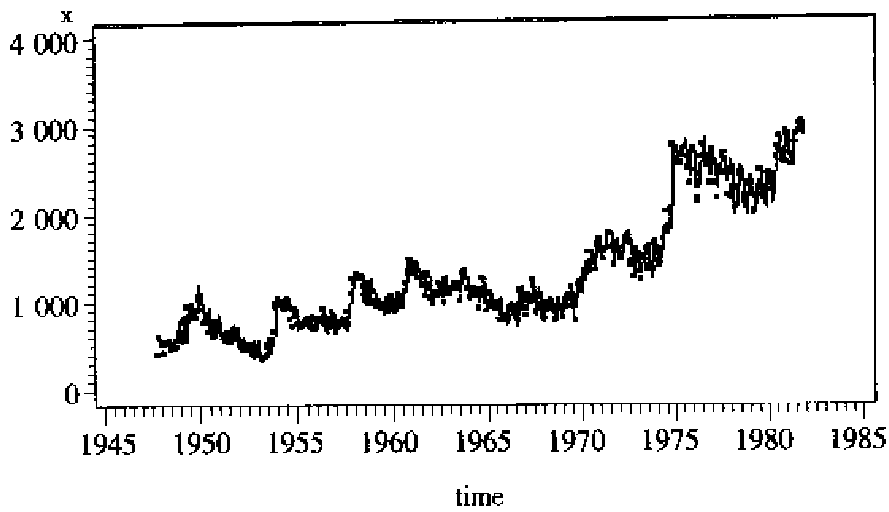


图 5—28　美国女性月度失业率序列拟合效果图

图中，点为序列观察值；曲线为序列拟合值。

5.3　Auto-Regressive 模型

ARIMA 模型的提出使人们对非平稳序列的拟合精度大大提高，自 1970 年 Box 和 Jenkins 提出这个模型以来，它已经成为最为经典的时间序列拟合模型。

但和传统的确定性因素分解方法比较，*ARIMA* 模型仍然有一些遗憾。它是使用差分方法提取确定性信息，差分方法的优点是对确定性信息的提取比较充分，它的缺点是很难对模型进行直观解释。所以当序列具有非常显著的确定性趋势或季节效应时，人们会怀念确定性因素分解方法对各种确定性效应的解释，但又因为它对残差信息的浪费而不敢轻易使用。

为了解决这个问题，人们构造了残差自回归（auto-regressive）模型。

5.3.1　模型结构

Auto-Regressive 模型的构造思想是首先通过确定性因素分解方法提取序列中主要的确定性信息：

$$x_t = T_t + S_t + \varepsilon_t \tag{5.4}$$

式中，T_t 为趋势效应拟合；S_t 为季节效应拟合。

考虑到因素分解方法对确定性信息的提取可能不够充分，因而需要进一步检

验残差序列 $\{\varepsilon_t\}$ 的自相关性。

如果检验结果显示残差序列自相关性不显著，说明确定性回归模型（5.4）对信息的提取比较充分，可以停止分析了。

如果检验结果显示残差序列自相关性显著，说明确定性回归模型（5.4）对信息的提取不充分，这时可以考虑对残差序列拟合自回归模型，进一步提取相关信息：

$$\varepsilon_t=\phi_1\varepsilon_{t-1}+\cdots+\phi_p\varepsilon_{t-p}+a_t$$

这样构造的模型：

$$x_t=T_t+S_t+\varepsilon_t$$

$$\varepsilon_t=\phi_1\varepsilon_{t-1}+\cdots+\phi_p\varepsilon_{t-p}+a_t$$

$$E(a_t)=0,\mathrm{Var}(a_t)=\sigma^2,\mathrm{Cov}(a_t,a_{t-i})=0,\forall i\geqslant 1$$

称为（残差）自回归模型（auto-regressive model）。

实践中，对趋势效应的拟合常用如下两种方式：

1. 自变量为时间 t 的幂函数

$$T_t=\beta_0+\beta_1\cdot t+\cdots+\beta_k\cdot t^k+\varepsilon_t$$

2. 自变量为历史观察值 $\{x_{t-1},x_{t-2},\cdots,x_{t-k}\}$

$$T_t=\beta_0+\beta_1\cdot x_{t-1}+\cdots+\beta_k\cdot x_{t-k}+\varepsilon_t$$

第二种方式和差分方式的原理雷同。

对季节效应的拟合也有两种常用的方式：

1. 给定季节指数

$$S_t=S_t'$$

式中，S_t' 是某个已知的季节指数。

2. 建立季节自回归模型

$$T_t=\alpha_0+\alpha_1\cdot x_{t-m}+\cdots+\alpha_l\cdot x_{t-lm}$$

假定固定周期为 m。

【例 5.6 续】 使用 Auto-Regressive 模型分析 1952—1988 年中国农业实际国民收入指数序列（数据见附录 1.14）。

图 5—10 显示该序列有显著的线性递增趋势，但没有季节效应，所以考虑拟建立如下结构的 Auto-Regressive 模型：

$$\begin{cases}x_t=T_t+\varepsilon_t,t=1,2,3,\cdots\\ \varepsilon_t=\phi_1\varepsilon_{t-1}+\cdots+\phi_p\varepsilon_{t-p}+a_t\\ E(a_t)=0,\mathrm{Var}(a_t)=\sigma^2,\mathrm{Cov}(a_t,a_{t-i})=0,\forall i\geqslant 1\end{cases}$$

对 T_t 分别尝试构造如下两个确定性趋势模型：

1. 变量为时间 t 的幂函数

$$T_t=\beta_0+\beta_1 \cdot t,\ t=1,2,3,\cdots$$

使用极大似然估计方法，得到该序列的趋势拟合模型为：

$$T_t=66.1491+4.5158t,\ t=1,2,3,\cdots$$

2. 变量为 1 阶延迟序列值 x_{t-1}

$$T_t=\beta_0+\beta_1 \cdot x_{t-1}$$

使用极大似然估计方法，并检验参数的显著性，得到该序列的趋势拟合模型为：

$$\hat{x}_t=1.0365x_{t-1},t=1,2,3,\cdots$$

这两个趋势拟合模型的拟合效果图如图 5—29 所示。

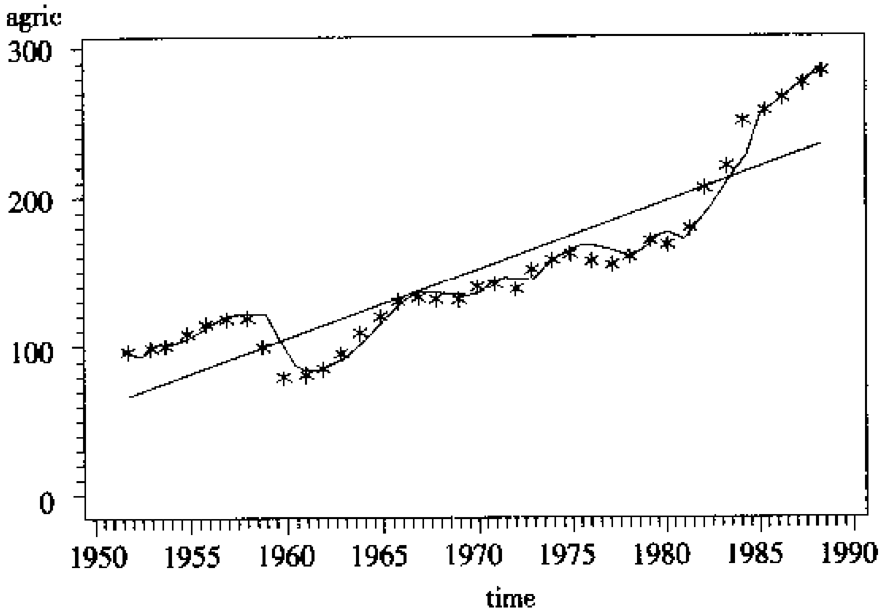


图 5—29 趋势拟合效果图

图中，星号为序列观察值；直线为 $T_t=66.1491+4.5158t$ 的拟合直线；曲线为 $T_t=1.0365x_{t-1}$的拟合曲线。

5.3.2 残差自相关检验

一、检验原理

确定性模型拟合好之后，我们要对该模型的拟合效果进行检验。

如果残差序列显示出纯随机的性质，即

$$E(\varepsilon_t,\varepsilon_{t-j})=0,\forall j\geqslant 1$$

就说明确定性模型拟合得非常好，已经能够充分提取序列中的相关信息了，我们不需要再对残差序列进行二次信息提取了，分析结束。

反之，如果残差序列显示出显著的自相关性，即

$$E(\varepsilon_t, \varepsilon_{t-j}) \neq 0, \exists j \geqslant 1$$

那就说明确定性模型拟合得不够精确，序列中的相关信息没有得到充分提取，我们应该对残差序列再次拟合，提取其中残存的相关信息，以提高模型拟合的精度。

二、Durbin-Waston 检验

Durbin-Waston 检验（简称 DW 检验）是由 J. Durbin 和 G. S. Watson 于 1950 年在考虑多元回归模型的残差独立性时提出的一个自相关检验统计量。我们把它借鉴过来进行时间序列残差自相关性检验。

下面以残差 1 阶自相关性检验为例介绍 DW 检验的原理。

原假设：残差序列不存在 1 阶自相关性，即

$$H_0: E(\varepsilon_t, \varepsilon_{t-1}) = 0 \Leftrightarrow H_0: \rho = 0$$

备择假设：残差序列存在 1 阶自相关性，即

$$H_1: E(\varepsilon_t, \varepsilon_{t-1}) \neq 0 \Leftrightarrow H_1: \rho \neq 0$$

构造 DW 检验统计量：

$$DW = \frac{\sum_{t=2}^{n} (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^{n} \varepsilon_t^2}$$

$$= \frac{\left(\sum_{t=2}^{n} \varepsilon_t^2 + \sum_{t=2}^{n} \varepsilon_{t-1}^2\right) - 2\sum_{t=2}^{n} \varepsilon_t \cdot \varepsilon_{t-1}}{\sum_{t=1}^{n} \varepsilon_t^2}$$

在大数场合，有

$$\sum_{t=2}^{n} \varepsilon_t^2 \approx \sum_{t=2}^{n} \varepsilon_{t-1}^2 \approx \sum_{t=1}^{n} \varepsilon_t^2$$

所以 DW 检验统计量近似等于

$$DW = 2\left(1 - \frac{\sum_{t=2}^{n} \varepsilon_t \cdot \varepsilon_{t-1}}{\sum_{t=1}^{n} \varepsilon_t^2}\right)$$

根据自相关系数的定义，有

$$\rho = \frac{\sum_{t=2}^{n} \varepsilon_t \cdot \varepsilon_{t-1}}{\sum_{t=1}^{n} \varepsilon_t^2}$$

即 $DW \cong 2(1-\rho)$

因为 $-1 \leqslant \rho \leqslant 1$，所以 $0 \leqslant DW \leqslant 4$。

当 $0 < \rho \leqslant 1$ 时，序列正相关，且 $\rho \to 1$ 时，$DW \to 0$；$\rho \to 0$ 时，$DW \to 2$。根据这个性质，可以确定两个临界值 d_L，d_U（$0 \leqslant d_L$，$d_U \leqslant 2$），当 $DW \leqslant d_L$ 时，序列显著正相关，当 $DW \leqslant d_U$ 时，序列显著不相关。当 $d_L \leqslant DW \leqslant d_U$ 时，现有数据还无法做出肯定判断，需要更多的信息支持。

当 $-1 < \rho \leqslant 0$ 时，序列负相关，且 $\rho \to -1$ 时，$DW \to 4$；$\rho \to 0$ 时，$DW \to 2$。同理可以确定两个临界值 d'_L，d'_U（$2 \leqslant d'_L$，$d'_U \leqslant 4$），当 $DW \leqslant d'_U$ 时，序列显著负相关，当 $DW \leqslant d'_L$ 时，序列显著不相关。当 $d'_L \leqslant DW \leqslant d'_U$ 时，现有数据还无法做出肯定判断，需要更多的信息支持。

根据 ρ 的对称性，有

$$\begin{cases} 2-d_U=d'_L-2 \\ 2-d'_L=d'_U-2 \end{cases} \Rightarrow \begin{cases} d'_L=4-d_U \\ d'_U=4-d_L \end{cases}$$

综上所述，我们可以得到如下相关性判断图示，如图 5—30 所示。

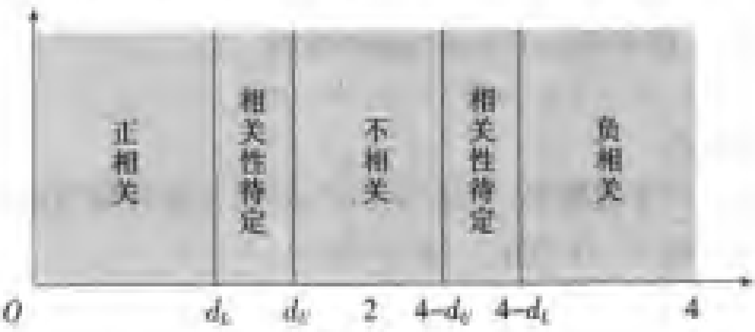


图 5—30

【例 5.6 续】 检验第一个确定性趋势模型：

$$x_t = 66.149\,1 + 4.515\,8t + \varepsilon_t,\ t = 1, 2, 3, \cdots$$

残差序列的自相关性。

检验结果如表 5—10 所示。（$\alpha = 0.05$）

表 5—10

检验统计量	统计量的值	d_L	d_U	P 值
DW	0.137 8	1.42	1.53	0.000 1

检验结果显示残差序列高度正自相关。为了充分提取相关信息，我们需要对残差序列进行再次拟合。

三、Durbin *h* 检验

值得注意的是，DW 统计量原本适用于回归模型残差自相关性的检验，它要

求自变量“独立”。Nerlove 和 Wallis 1966 年指出，在自回归场合，即当回归因子包含延迟因变量（计量经济学称之为内生变量）时，有

$$x_t=\beta_0+\beta_1\cdot x_{t-1}+\cdots+\beta_k\cdot x_{t-k}+\varepsilon_t$$

残差序列 $\{\varepsilon_t\}$ 的 DW 统计量是一个有偏统计量，当 ρ 趋向于零时，$d\neq 2$。在这种场合下使用 DW 统计量容易产生残差序列正自相关性不显著的误判。

基于这个原因，在 *ARIMA* 模型场合，我们都是使用 Q 统计量和 LB 统计量检验残差序列的自相关性。

为了克服 DW 检验的有偏性，Durbin 在 1970 年提出了 DW 统计量的两个修正统计量：Durbin t 和 Durbin h 统计量，这两个统计量渐近等价。Durbin h 统计量为：

$$D_h=DW\frac{n}{1-n\sigma_\beta^2}$$

式中，n 为观察值序列长度；σ_β^2 为延迟因变量系数的最小二乘估计的方差。

通过对 DW 统计量的修正，Durbin t 和 Durbin h 统计量有效地提高了检验的精度，代替 DW 检验统计量，成为延迟因变量场合常用的自相关检验统计量。

【例 5.6 续】 检验第二个确定性趋势模型

$$x_t=1.036\ 5x_{t-1}+\varepsilon_t,\quad t=1,\ 2,\ 3,\ \cdots$$

残差序列的自相关性。

这是一个延迟因变量模型，所以采用 Durbin h 统计量检验残差序列的自相关性。检验结果如表 5—11 所示。($\alpha=0.05$)

表 5—11

检验统计量	统计量的值	P 值
Durbin h	2.803 8	0.002 5

注：Durbin h 统计量没有临界值表，它的 P 值由 SAS 软件算出。

检验结果显示残差序列高度自相关。为了充分提取相关信息，我们需要对残差序列进行再次拟合。

5.3.3 模型拟合

一、确定自回归模型的阶数

【例 5.6 续】 以第一个确定性模型为例，残差序列值为：

$$\varepsilon_t=x_t-T_t=x_t-66.149\ 1-4.515\ 8,\quad t=1,\ 2,\ \cdots$$

残差序列的自相关图显示出典型的短期相关性，如图 5—31 所示。

```
                         Autocorrelations
Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  1.000 00    |                    |********************|
  0.860 45    |                .   |*****************   |
  0.655 06    |              .     |*************       |
  0.440 82    |            .       |*********   .       |
  0.215 16    |           .        |****         .      |
  0.012 76    |           .        |             .      |
  -.123 60    |           .      **|             .      |
  -.210 01    |           .    ****|             .      |
  -.203 05    |           .    ****|             .      |
  -.159 35    |           .     ***|             .      |

"."marks two standard errors
```

图 5—31　残差序列自相关图

偏自相关图呈现典型的 2 阶截尾性，如图 5—32 所示。

```
Partial Autocorrelations

-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
|                .   |*****************   |
|          *******|  .                |
|               **|  .                |
|          .  ****|  .                |
|          .     *|  .                |
|          .      |* .                |
|          .     *|  .                |
```

图 5—32　残差序列偏自相关图

所以尝试对残差序列 $\{\varepsilon_t\}$ 拟合 $AR(2)$ 模型：

$$\varepsilon_t = \phi_1\varepsilon_{t-1} - \phi_2\varepsilon_{t-2} + a_t$$

相同方法分析，可以确定第二个模型残差序列的自相关阶数为 1。

二、参数估计

尽管前面已经确定了原序列回归模型的口径，相应地，可以确定残差序列的值。单纯根据残差序列的值可以非常容易地确定残差自回归模型的口径：

$$\varepsilon_t = \hat{\phi}_1\varepsilon_{t-1} + \cdots + \hat{\phi}_p\varepsilon_{t-p}$$

但是，由于残差序列 $\{\varepsilon_t\}$ 与序列回归值 $\{\hat{x}_t\}$ 之间具有相关性，所以按照上述方法将它们分开求解实际上忽略了相关性的影响，这将会降低模型拟合的精度。

所以，参数最优估计应该是在前面分析的基础上，确定回归模型的结构和残差自回归模型的阶数，将所有参数联合求解。

【例 5.6 续】 使用极大似然估计，联立求解，确定 Auto-Regressive 模型的口径。

参数估计结果如下：

(1) 模型一

$$\begin{cases} x_t = 69.149\ 1 + 4.515\ 8t + \varepsilon_t \\ \varepsilon_t = 1.485\ 9\varepsilon_{t-1} - 0.584\ 8\varepsilon_{t-2} + a_t \end{cases}$$

式中，$\{a_t\}$ 为零均值白噪声序列。

(2) 自回归模型二

$$\begin{cases} x_t = 1.033x_{t-1} + \varepsilon_t \\ \varepsilon_t = 0.461\ 5\varepsilon_{t-1} + a_t \end{cases}$$

式中，$\{a_t\}$ 为零均值白噪声序列。

分析至此，我们一共为 1952—1988 年中国农业实际国民收入指数序列拟合了三个模型，比较我们拟合的这三个模型的优劣。如表 5—12 所示。

表 5—12

模型	AIC	SBC
$ARIMA(0,1,1)$模型 $(1-B)x_t = 4.996\ 61 + (1+0.707\ 66B)\varepsilon_t$	249.330 5	252.497 6
Auto-Regressive 模型一 $\begin{cases} x_t = 69.149\ 1 + 4.515\ 8t + \varepsilon_t \\ \varepsilon_t = 1.485\ 9\varepsilon_{t-1} - 0.584\ 8\varepsilon_{t-2} + a_t \end{cases}$	260.845 4	267.289 1
Auto-Regressive 模型二 $\begin{cases} x_t = 1.033x_{t-1} + \varepsilon_t \\ \varepsilon_t = 0.461\ 5\varepsilon_{t-1} + a_t \end{cases}$	250.631 7	253.798 7

比较结果显示，$ARIMA$ (0，1，1) 模型拟合得最好，其次为以延迟因变量为回归因子的 Auto-Regressive 模型二，拟合效果最差的是以时间变量为回归因子的 Auto-Regressive 模型一。究其原因主要是时间自变量模型对确定性信息的提取精度差了一点。

而从直观解释的角度考虑 Auto-Regressive 模型一：

$$\begin{cases} x_t = 69.149\ 1 + 4.515\ 8t + \varepsilon_t \\ \varepsilon_t = 1.485\ 9\varepsilon_{t-1} - 0.584\ 8\varepsilon_{t-2} + a_t \end{cases}$$

更易于直观解释原序列的波动规律：中国农业实际国民收入指数拥有一个长期的线性递增趋势，平均每期增长4.515 8。同时它还受到诸多随机因素的影响，导致随机波动序列具有短期自相关性。

5.4 异方差的性质

1982 年 Engle 在分析英国通货膨胀率序列时，发现经典的 *ARIMA* 模型始终无法取得理想的拟合效果。为什么会这样呢？经过对残差序列仔细的研究，他发现问题出在残差序列具有异方差性。

5.4.1 异方差的影响

使用 *ARIMA* 模型拟合非平稳序列时，对残差序列有一个重要假定——残差序列 $\{\varepsilon_t\}$ 为零均值白噪声序列。换言之，残差序列要满足如下三个假定条件。

一、零均值

$E(\varepsilon_t)=0$

二、纯随机

$\mathrm{Cov}(\varepsilon_t,\varepsilon_{t-i})=0,\forall i\geqslant 1$

三、方差齐性

$\mathrm{Var}(\varepsilon_t)=\sigma_\varepsilon^2$

如果方差齐性假定不成立，即随机误差序列的方差不再是常数了，它会随着时间的变化而变化，可以表示为时间的某个函数：

$\mathrm{Var}(\varepsilon_t)=h(t)$

这种情况被称作异方差。

在残差序列的这三个假定中，零均值假定最容易实现，只要对序列进行中心化处理就可以实现。所以这个假定通常无须检验。

纯随机假定一直是我们重点监控的对象。如果这个假定不满足就说明残差序列中还蕴含着使得提取的自相关信息。为了有效检验这个假定条件是否成立，统计学家们构造了许多适用于不同场合的自相关检验统计量，比如我们前面介绍过的 Q 统计量、LB 统计量、DW 统计量等等。

只有第三个假定——方差齐性假定，在此之前我们没有进行任何检验。在缺省检验的情况下就默认残差序列一定满足这个条件。但实际上，这个假定条件并不总是满足。忽视异方差的存在会导致残差的方差被严重低估，继而参数显著性检验容易犯纳伪错误，这使得参数的显著性检验失去意义，最终导致模型的拟合精度受影响。所以为了提高模型拟合的精度，我们需要对残差序列进行方差齐性

检验，并且对异方差序列进行深入分析。

5.4.2 异方差的直观诊断

一、残差图

当残差序列 $\{\varepsilon_t\}$ 方差齐性时，它应该在零值附近随机波动，不带任何趋势，否则就显示出异方差的性质了，如图 5—33、图 5—34、图 5—35、图 5—36 所示。

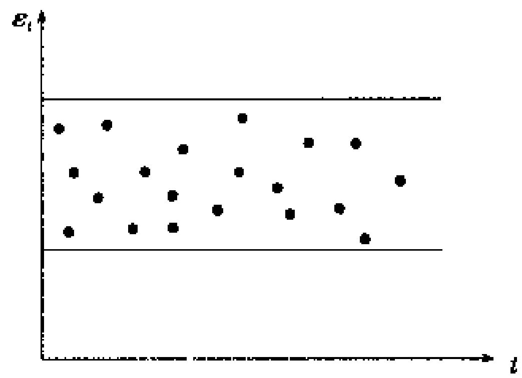


图 5—33 方差齐性残差图

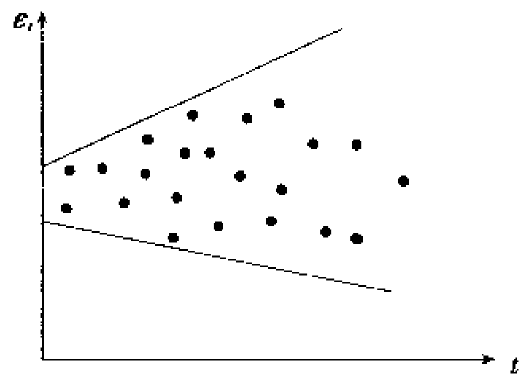


图 5—34 递增型异方差

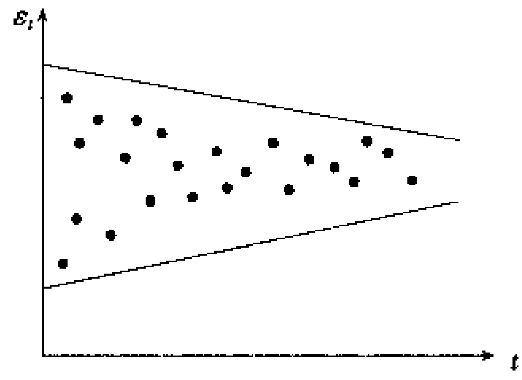


图 5—35 递减型异方差

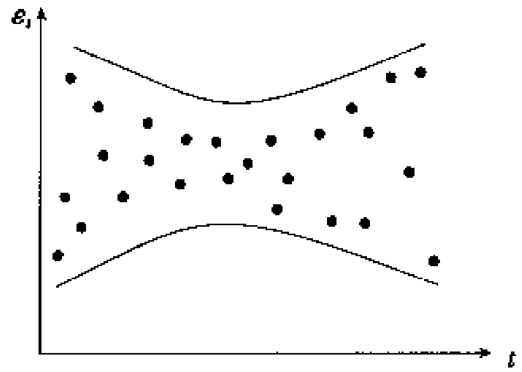


图 5—36 综合型异方差

二、残差平方图

由于残差序列的方差实际上就是它平方的期望，即

$$\mathrm{Var}(\varepsilon_t)=E(\varepsilon_t^2)$$

所以残差序列是否方差齐性，主要是考察 ε_t^2 的性质。我们可以借助残差平方图——ε_t^2 关于 t 变化的二维坐标图，对残差序列的方差齐性进行直观诊断。

和残差图的判断原则一样，假设方差齐性满足的话，有

$$E(\varepsilon_t^2)=\sigma_\varepsilon^2$$

这意味着 ε_t^2 应该在某个常数值 σ_ε^2 附近随机波动，它不应该具有任何明显的趋势，否则就呈现出异方差性。

【例 5.11】 直观考察美国 1963 年 4 月—1971 年 7 月短期国库券的月度收益率序列的方差齐性（数据见附录 1.18）。

该序列时序图显示序列显著非平稳，如图 5—37 所示。

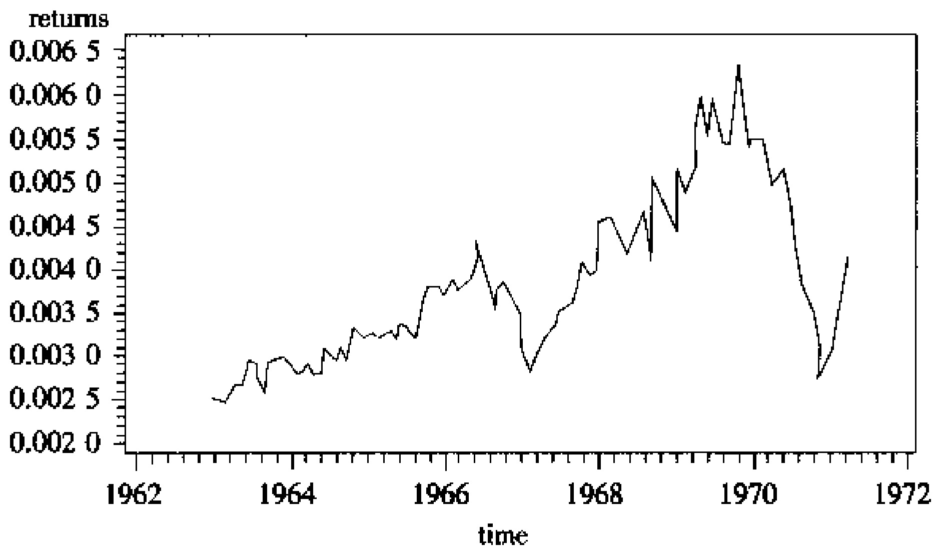


图 5—37 短期国库券月度收益率时序图

对原序列进行1阶差分，差分后残差序列显示出均值平稳，但方差递增的性质，如图5—38所示。

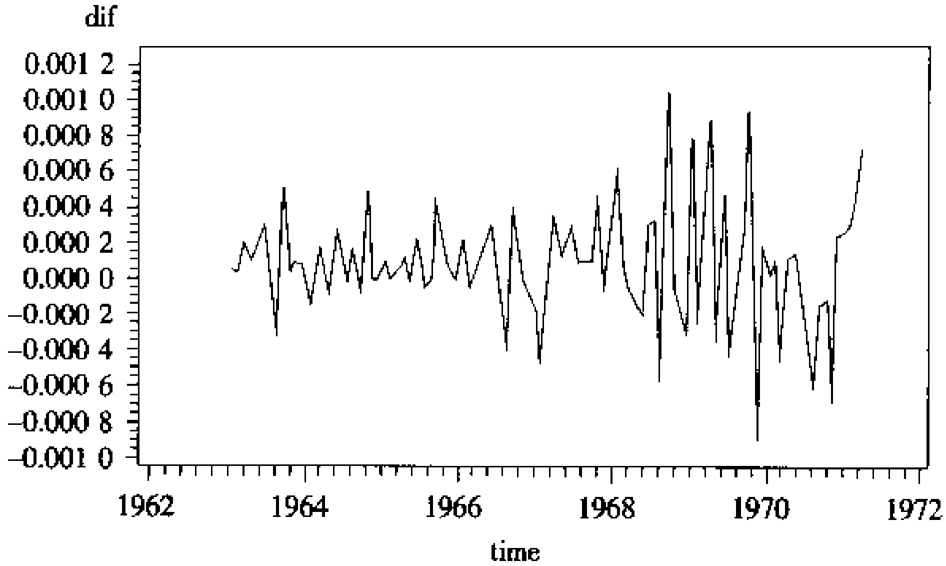


图5—38 1阶差分后残差序列时序图

观察1阶差分后序列残差平方图，可以看出残差平方具有显著的差异性，同样得出残差序列异方差的结论。如图5—39所示。

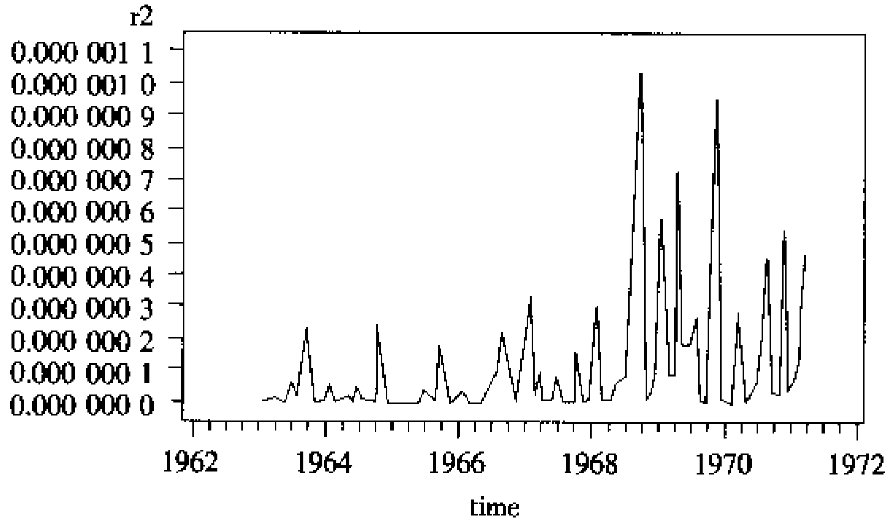


图5—39 1阶差分后残差平方图

当残差序列异方差时，我们需要对它进行进一步的处理，处理思路有两种：

1. 假如已知异方差函数具体形式，进行方差齐性变化。

2. 假如不知异方差函数的具体形式，拟合条件异方差模型。

5.5 方差齐性变换

一、使用场合

假设序列显示出显著的异方差性，且方差 σ_t^2 与均值 μ_t 之间具有某种函数关系：

$$\sigma_t^2=h(\mu_t)$$

式中，$h(\cdot)$是某个已知函数。

在这种场合下,我们的处理思路是尝试寻找一个转换函数 $g(\cdot)$，使得经转换后的变量 $g(x_t)$满足方差齐性：

$$\mathrm{Var}[g(x_t)]=\sigma^2$$

二、转换函数的确定

将 $g(x_t)$在 μ_t 附近作 1 阶泰勒展开：

$$g(x_t)\cong g(\mu_t)+(x_t-\mu_t)g'(\mu_t)$$

式中，$g'(\cdot)$为 $g(\cdot)$的 1 阶导数。则 $g(x_t)$的方差近似等于

$$\begin{aligned}\mathrm{Var}[g(x_t)]&\cong\mathrm{Var}[g(\mu_t)+(x_t-\mu_t)g'(\mu_t)]\\&=[g'(\mu_t)]^2\cdot\mathrm{Var}(x_t)\\&=[g'(\mu_t)]^2h(\mu_t)\end{aligned}$$

显然，要使得 $\mathrm{Var}[g(x_t)]$等于常数，转换函数 $g(\cdot)$必须与$\sqrt{h(\cdot)}$具有倒函数关系：

$$g'(\mu_t)=\frac{1}{\sqrt{h(\mu_t)}}$$

在实践中，许多金融时间序列都呈现出异方差的性质，而且通常序列的标准差与其水平之间具有某种正比关系，即序列的水平低时，序列的波动范围小，序列的水平高时，序列的波动范围大。对于这种异方差的性质，最简单的假定为：

$$\sigma_t=\mu_t$$

即等价于

$$h(\mu_t)=\mu_t^2$$

要使得原序列经过适当转换后方差齐性，就必须满足

$$g'(\mu_t)=\frac{1}{\sqrt{h(\mu_t)}}=\frac{1}{\mu_t}$$

等价推导出

$$g(\mu_t)=\log(\mu_t)$$

这意味着对于标准差与水平成正比关系的异方差序列，对数变换可以有效地实现方差齐性。

【例 5.11 续】 对美国 1963 年 4 月—1971 年 7 月短期国库券的月度收益率序列使用方差齐性变换方法进行分析。

在前面的分析中我们看到该序列残差的标准差与序列值之间具有一定的正相关性，假设它们之间具有正比关系：

$$\sigma_t=x_t$$

对原序列进行对数变换：

$$y_t=\log(x_t)$$

对数序列时序图显示它保持了原序列的变化趋势，如图 5—40 所示。

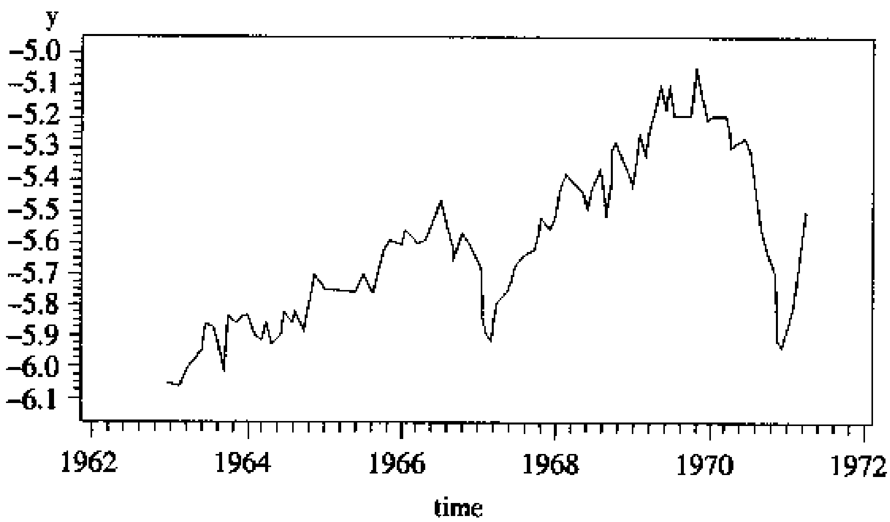


图 5—40 对数序列时序图

对它进行 1 阶差分：

$$\nabla y_t=\log(x_t)-\log(x_{t-1})$$

差分后的残差图如图 5—41 所示。

残差图直观显示残差序列波动平稳，白噪声检验显示该残差序列纯随机，如表 5—13 所示。

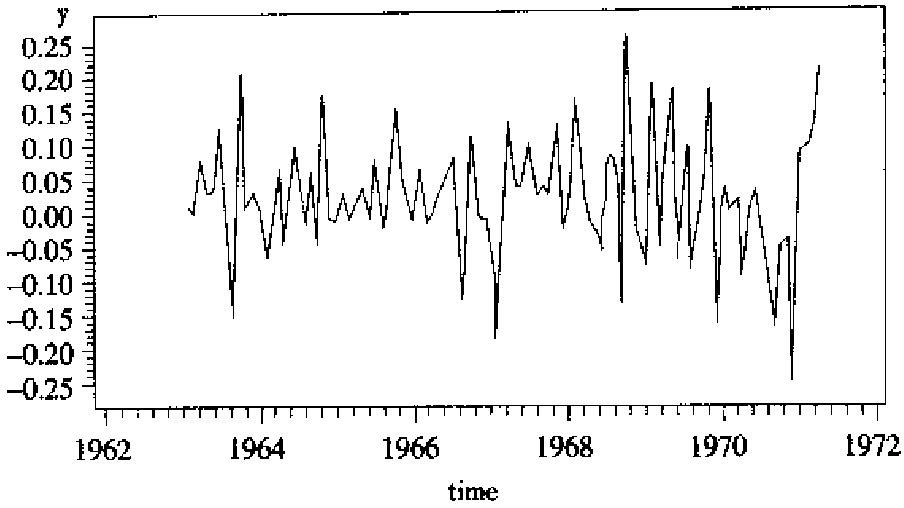


图 5—41　对数序列 1 阶差分后残差图

表 5—13

延迟阶数	LB统计量	P值
6	3.58	0.733 7
12	10.82	0.544 1
18	21.71	0.245 2

这说明，通过方差齐性变换，得到该序列的拟合模型为：

$$\nabla \log(x_t)=\varepsilon_t$$

式中，$\{\varepsilon_t\}$ 为零均值白噪声序列。

该模型拟合效果如图 5—42 所示。

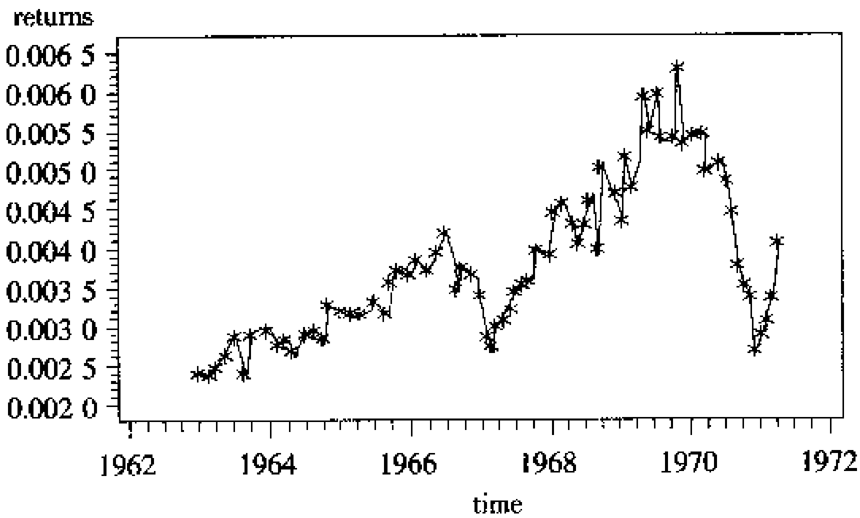


图 5—42　拟合效果图

图中，星号为序列观察值；曲线为序列拟合值。从拟合图中可以看出本例使用方差齐性变换的方法进行分析，拟合效果不错。

5.6 条件异方差模型

方差齐性变换为异方差序列的精确拟合提供了一个很好的解决方法，但这个方法更多的是具有理论上的意义，在实践中的可操作性并不强。因为要使用方差齐性变换必须得事先知道异方差函数的形式，而这通常是不可能的。

实践中，我们只能根据残差图及残差平方图所显示出来的特点，使用一些常用的函数形式估计异方差函数。这体现在进行金融时间序列分析时，由于金融序列的标准差与水平之间通常具有某种正相关关系，所以异方差函数经常被假定为：

$$h(\mu_t)=\mu_t^2$$

继而导致对数变换在进行金融时间序列分析时被普遍采用。但大量的实践证明这种假定太简单化了，对数变换通常只能使绝大多数金融时间序列的异方差程度得到改善，但无法真正实现方差齐性。

为了更精确地估计异方差函数，Engle 于 1982 年提出了条件异方差模型。

5.6.1 模型结构

一、*ARCH* 模型

ARCH 模型的全称是自回归条件异方差模型（autoregressive conditional heteroskedastic），它的构造原理如下：

假设在历史数据已知的情况下，零均值、纯随机残差序列具有异方差性：

$$\mathrm{Var}(\varepsilon_t)=h_t$$

在正态分布的假定下，有

$$\varepsilon_t/\sqrt{h_t}\sim N(0,1)$$

异方差等价于残差平方的均值：

$$E(\varepsilon_t^2)=h_t$$

使用残差平方序列的自相关系数，可以考察异方差函数的自相关性：

$$\rho_k=\sqrt{\frac{\mathrm{Cov}(\varepsilon_t^2,\varepsilon_{t-k}^2)}{\mathrm{Var}(\varepsilon_t^2)}}$$

考察的结果无外乎如下两种：

1. 自相关系数恒为零

$$\rho_k=0,\ k=1,\ 2,\ \cdots$$

这说明异方差函数纯随机。此时，历史数据对未来异方差的估计一点作用都没有。这是最难分析的一种情况，至今没有有效的方法提取其中的异方差信息。

2. 存在某个自相关系数不为零

$$\rho_k\neq 0,\ k\geqslant 1$$

这是我们主要分析的情况。

误差平方序列的自相关系数不恒为零，说明异方差函数存在自相关性，这使得我们有可能通过构造残差平方序列的自回归模型来拟合异方差函数：

$$h_t=E(\varepsilon_t^2)=\omega+\sum_{j=1}^{q}\lambda_j\varepsilon_{t-j}^2$$

这样构造的模型称为 q 阶自回归条件异方差模型，简记为 $ARCH(q)$，它的完整结构为：

$$\begin{cases}x_t=f(t,x_{t-1},x_{t-2},\cdots)+\varepsilon_t\\ \varepsilon_t=\sqrt{h_t}e_t\\ h_t=\omega+\sum_{j=1}^{q}\lambda_j\varepsilon_{t-j}^2\end{cases}$$

式中，$f(t,x_{t-1},x_{t-2},\cdots)$ 为 $\{x_t\}$ 的 Auto-Regressive 模型；$e_t\overset{i,i,d}{\sim}N(0,1)$。

二、*GARCH* 模型

ARCH 模型的实质是使用误差平方序列的 q 阶移动平均拟合当期异方差函数值。由于移动平均模型具有自相关系数 q 阶截尾性，所以 *ARCH* 模型实际上只适用于异方差函数短期自相关过程。

但是在实践中，有些残差序列的异方差函数是具有长期自相关性的，这时如果使用 *ARCH* 模型拟合异方差函数，将会产生很高的移动平均阶数，这会增加参数估计的难度并最终影响 *ARCH* 模型的拟合精度。

为了修正这个问题，Bollerslov 在 1985 年提出了广义自回归条件异方差（generalized autoregressive conditional heteroskedastic）模型，它的结构如下：

$$\begin{cases}x_t=f(t,x_{t-1},x_{t-2},\cdots)+\varepsilon_t\\ \varepsilon_t=\sqrt{h_t}e_t\\ h_t=\omega+\sum_{i=1}^{p}\eta_i h_{t-i}+\sum_{j=1}^{q}\lambda_j\varepsilon_{t-j}^2\end{cases}$$

式中，$f(t,x_{t-1},x_{t-2},\cdots)$ 为 $\{x_t\}$ 的回归函数；$e_t\overset{i,i,d}{\sim}N(0,1)$。这个模型简记为 $GARCH(p,q)$。

GARCH 模型实际上就是在 *ARCH* 模型的基础上，增加考虑了异方差函数

的 p 阶自相关性。它可以有效地拟合具有长期记忆性的异方差函数。

显然 *ARCH* 模型是 *GARCH* 模型的一种特例，*ARCH*(q) 模型实际上就是 $p=0$ 的模型 *CARCH*（p，q）。

三、*GARCH* 模型的变体

GARCH 模型是至今为止最常用、最可行的异方差序列拟合模型。但它的有效使用必须满足如下两个约束条件。

条件 1：参数非负

$$\omega>0,\ \eta_i\geqslant 0,\ \lambda_j\geqslant 0$$

条件 2：参数有界

$$\sum_{j=1}^{p}\eta_i+\sum_{j=1}^{q}\lambda_j<1$$

这两个约束条件限制了 *GARCH* 模型的使用面。为了拓宽 *GARCH* 模型的使用，许多统计学家从不同的角度出发，构造了多个 *GARCH* 模型的变体。

1. 指数 *GARCH* 模型（*EGARCH*）

为了放宽 *GARCH* 模型中参数非负的约束，Nelson 于 1991 年提出了 *EGARCH* 模型，该模型结构如下：

$$\begin{cases}x_t=f(t,x_{t-1},x_{t-2},\cdots)+\varepsilon_t\\ \varepsilon_t=\sqrt{h_t}e_t\\ \ln(h_t)=\omega+\sum_{i=1}^{p}\eta_i\ln(h_{t-i})+\sum_{j=1}^{q}\lambda_j g(e_t)\\ g(e_t)=\theta e_t+\gamma[\mid e_t\mid-E\mid e_t\mid]\end{cases}$$

式中，$f(t,x_{t-1},x_{t-2},\cdots)$ 为 $\{x_t\}$ 的回归函数；$e_t\overset{i,i,d}{\sim}N(0,1)$。

这样构造的 *EGARCH* 模型不再需要参数非负约束。

2. 方差无穷 *GARCH* 模型（*IGARCH*）

当 *GARCH* 模型平稳时，ε_t 的无条件方差为：

$$\mathrm{Var}(\varepsilon_t)=\frac{\omega}{1-(\sum_{i=1}^{p}\eta_i+\sum_{j=1}^{q}\lambda_j)}$$

因而 *GARCH* 模型参数有界约束的目的是为了保证 ε_t 的无条件方差有界，*GARCH* 模型能实现宽平稳。

如果把这个约束条件改写为：

$$\sum_{i=1}^{p}\eta_i+\sum_{j=1}^{q}\lambda_j=1$$

就构成 IGARCH 模型。

对 IGARCH 模型而言，ε_t 的无条件方差无界，所以它不是宽平稳的，但它却是严平稳的。

3. 依均值 GARCH 模型（GARCH-M）

有时序列均值与条件方差之间具有某种相关关系，这时可以把条件标准差作为附加回归因子建模，模型结构如下：

$$\begin{cases} x_t = f(t, x_{t-1}, x_{t-2}, \cdots) + \delta\sqrt{h_t} + \varepsilon_t \\ \varepsilon_t = \sqrt{h_t}e_t \\ h_t = \omega + \sum_{i=1}^{q}\eta_i h_{t-i} + \sum_{j=1}^{q}\lambda_j\varepsilon_{t-j}^2 \end{cases}$$

式中，$f(t, x_{t-1}, x_{t-2}, \cdots)$ 为 $\{x_t\}$ 的回归函数；$e_t \overset{i.i.d}{\sim} N(0,1)$。

四、AR-GARCH 模型

对残差序列 $\{\varepsilon_t\}$ 拟合 GARCH 模型有一个基本要求：$\{\varepsilon_t\}$ 为零均值，纯随机、异方差序列。有时回归函数 $f(t, x_{t-1}, x_{t-2}, \cdots)$ 不能充分提取原序列 $\{\varepsilon_t\}$ 中的相关信息，残差序列可能具有自相关性，而不是纯随机的。这时需要对 $\{\varepsilon_t\}$ 先拟合自回归模型，再考察自回归残差序列 $\{\upsilon_t\}$ 的方差齐性，如果 $\{\upsilon_t\}$ 异方差，对它拟合 GARCH 模型。这样构造的模型称为 $AR(m)-GARCH(p, q)$ 模型：

$$\begin{cases} x_t = f(t, x_{t-1}, x_{t-2}, \cdots) + \varepsilon_t \\ \varepsilon_t = \sum_{k=1}^{m}\beta_k\varepsilon_{t-k} + \upsilon_t \\ \upsilon_t = \sqrt{h_t}e_t \\ h_t = \omega + \sum_{i=1}^{p}\eta_i h_{t-i} + \sum_{j=1}^{q}\lambda_j\upsilon_{t-j}^2 \end{cases}$$

式中，$f(t, x_{t-1}, x_{t-2}, \cdots)$ 为 $\{x_t\}$ 的回归函数；$e_t \overset{i.i.d}{\sim} N(0,1)$。

5.6.2 模型拟合

GARCH 模型的拟合遵循如下六个步骤：

一、回归拟合

GARCH 模型拟合的第一步是根据观察值序列的性质，拟合回归模型 $f(t, x_{t-1}, x_{t-2}, \cdots)$。

【例 5.12】 某金融时间序列的时序图如图 5—43 所示（数据见附录 1.19）。

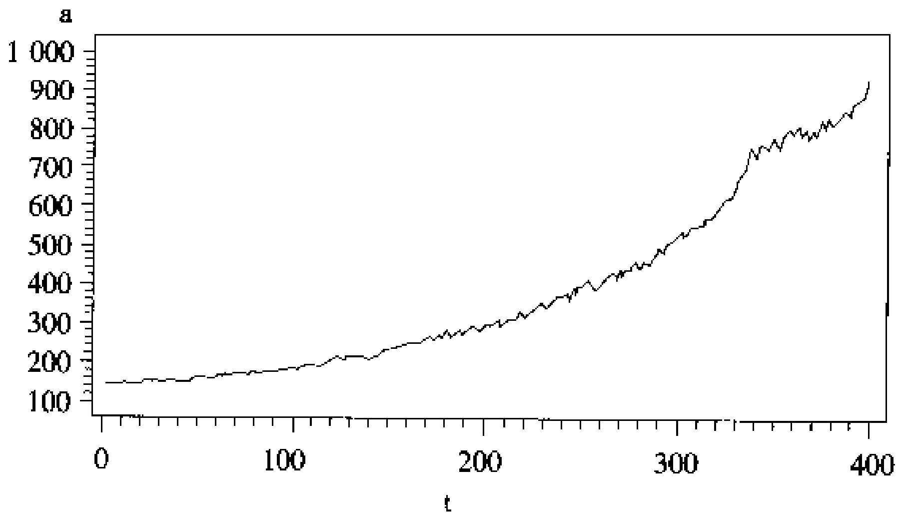


图 5—43　金融时间序列时序图

时序图显示序列具有显著的递增趋势，这是典型的自相关特征，尝试构造线性自相关函数：

$x_t = \alpha_1 x_{t-1} + \varepsilon_t$

使用极大似然估计得到未知参数的估计使：

$\hat{\alpha}_1 = 1.0053$

参数的显著性检验显示参数高度显著。

二、残差自相关性检验

观察值减去回归拟合值即为残差：

$\varepsilon_t = x_t - \hat{\beta}_1 x_{t-1}$

通过 DW 统计量检验残差序列的自相关性。

【例 5.12 续】　对残差序列进行自相关性检验，检验结果显示：

Durbin $h = -2.6011$，$\Pr(Dh < -2.6011) < 0.0046$

自相关性检验显示残差显著自相关。因此，还需要对残差序列拟合自回归模型：

$$\varepsilon_t = \sum_{k=1}^{m} \beta_k \varepsilon_{t-k} + \upsilon_t$$

使用逐步回归的方法，得到残差自回归模型如下：

$\varepsilon_t = -0.1559\varepsilon_{t-1} - 0.407\varepsilon_{t-2} + \upsilon_t$

由此产生新的残差序列 $\{\upsilon_t\}$。

三、异方差自相关性检验

经过前面两个步骤的处理，残差序列 $\{\upsilon_t\}$ 基本实现了零均值、纯随机的性质，现在需要考察它的异方差性。而异方差检验的实质是进行异方差相关性检

验。最常用的两个异方差检验是：Portmantea Q 检验和 LM 检验。

（一）Portmantea Q 检验

1983 年 Mcleod 和 Li 在 LB 统计量的基础上提出了 Portmanteau Q 统计量，用于检验残差平方序列的自相关性。

1. 假设条件

H_0：残差平方序列纯随机↔H_1：残差平方序列具有自相关性

用自相关系数表示，该假设条件等价于

$H_0:\rho_1=\rho_2=\cdots=\rho_q=0 \leftrightarrow H_1:\rho_k,\rho_2,\cdots,\rho_q$ 不全为零

2. 检验统计量

$$Q(q)=n(n+2)\sum_{i=1}^{q}\frac{\rho_i^2}{n-i}$$

式中，n 为观察序列长度；ρ_i 为残差序列延迟 i 自相关系数，记 $\hat{\sigma}^2=\frac{\sum_{t=1}^{n}\varepsilon_t^2}{n}$。

$$\rho_i=\sqrt{\frac{\sum_{t=i+1}^{n}(\varepsilon_t^2-\hat{\sigma}^2)(\varepsilon_{t-i}^2-\hat{\sigma}^2)}{\sum_{t=1}^{n}(\varepsilon_t^2-\hat{\sigma}^2)^2}}$$

Portmanteau Q 统计量近似服从自由度为 $q-1$ 的 χ^2 分布。

$$Q(q)\sim\chi^2(q-1)$$

3. 检验结果

检验结果如表 5 14 所示。

表 5—14

显著性水平	临界值	Portmanteau Q 统计量	P 值	检验结果
α	$\chi^2_{1-\alpha}(q-1)$	$Q(q)\geqslant\chi^2_{1-\alpha}(q-1)$	$\leqslant\alpha$	拒绝 H_0
		$Q(q)\geqslant\chi^2_{1-\alpha}(q-1)$	$>\alpha$	接受 H_0

（二）拉格朗日乘子检验

1982 年 Engle 为了确定 $ARCH(q)$ 模型的阶，提出了拉格朗日乘子检验（Lagrange multiplier test），简记为 LM 检验。

1. 假设条件

H_0：残差平方序列纯随机↔H_1：残差平方序列具有自相关性

等价于

$H_0:\rho_1=\rho_2=\cdots=\rho_q=0 \leftrightarrow H_1:\rho_k,\rho_2,\cdots,\rho_q$ 不全为零

2. 检验统计量

$$LM(q)=W'W$$

式中：

$$W=\left(\frac{\rho_1^2}{\hat{\sigma}^2},\frac{\rho_2^2}{\hat{\sigma}^2},\cdots,\frac{\rho_q^2}{\hat{\sigma}^2}\right)'$$

$$\rho_i=\sqrt{\frac{\sum_{t=i+1}^{n}(\varepsilon_t^2-\hat{\sigma}^2)(\varepsilon_{t-i}^2-\hat{\sigma}^2)}{\sum_{t=1}^{n}(\varepsilon_t^2-\hat{\sigma}^2)^2}}$$

$$\hat{\sigma}^2=\frac{\sum_{t=1}^{n}\varepsilon_t^2}{n}。$$

$LM(q)$统计量近似服从自由度为 $q-1$ 的 χ^2 分布。

$$LM(q)\sim\chi^2(q-1)$$

3. 检验结果

检验结果如表 5—15 所示。

表 5—15

显著性水平	临界值	LM(q)统计量	P 值	检验结果
α	$\chi^2_{1-\alpha}(q-1)$	$LM(q)\geqslant\chi^2_{1-\alpha}(q-1)$	$\leqslant\alpha$	拒绝 H_0
		$LM(q)<\chi^2_{1-\alpha}(q-1)$	$>\alpha$	接受 H_0

【5.12 续】 对残差序列 $\{\upsilon_t\}$ 进行异方差检验，检验结果如表 5—16 所示。

表 5—16

阶数	Portmantea Q 统计量	P 值	LM 统计量	P 值
1	21.045 7	<0.000 1	20.9461	<0.000 1
2	69.583 6	<0.000 1	57.621 9	<0.000 1
3	91.206 2	<0.000 1	62.896 3	<0.000 1
4	103.772 5	<0.000 1	63.087 4	<0.000 1
5	105.121 6	<0.000 1	65.926 8	<0.000 1
6	105.175 3	<0.000 1	68.550 4	<0.000 1
7	105.485 8	<0.000 1	68.680 5	<0.000 1
8	116.660 5	<0.000 1	86.644 6	<0.000 1
9	131.200 3	<0.000 1	102.367 3	<0.000 1
10	185.995 7	<0.000 1	135.577 4	<0.000 1
11	205.330 4	<0.000 1	135.646 5	<0.000 1
12	417.184 4	<0.000 1	230.353 7	<0.000 1

Portmantea Q 统计量和 LM 统计量的检验结果都显示残差序列 $\{\upsilon_t\}$ 具有显

著的异方差自相关性，而且是长期自相关性。假设异方差函数为 h_t，则有

$$v_t/\sqrt{h_t} \sim N(0,1)$$

四、*ARCH* 模型阶数的估计

由于 Portmantea Q 统计量和 LM 统计量都显示出残差序列 $\{v_t\}$ 具有异方差长期自相关性，所以不妨尝试拟合 $GARCH(1,1)$模型

$$h_t=\omega+\eta_1 h_{t-1}+\lambda_1 v_{t-1}^2$$

五、参数估计

使用极大似然估计方法，采用迭代技术，计算所有未知参数的估计值。

【例 5.12 续】 本例实际上最终拟合的是 $AR(2)-GARCH(1,1)$模型，通过极大似然估计确定模型的口径为：

$$\begin{cases} x_t=1.0046x_{t-1}+\varepsilon_t \\ \varepsilon_t=-0.1559\varepsilon_{t-1}-0.407\varepsilon_{t-2}+v_t \\ v_t=\sqrt{h_t}e_t \\ h_t=0.0951+0.8999h_{t-1}+0.1053v_{t-1}^2 \end{cases}$$

参数显著性检验显示，除了 $\omega=0.0951$ 不显著外，其他参数均显著。

六、正态性检验

模型拟合好之后，我们要对模型的有效性进行检验。我们采用的检验工具是正态性检验。正态性检验的原理基于 $GARCH$ 模型的基本假定之一：当异方差模型拟合好之后，如果拟合得准确，函数 $\varepsilon_t/\sqrt{h_t}$（如果是 $AR-GARCH$ 模型则为函数 $v_t/\sqrt{h_t}$）会呈现出显著的标准正态分布特征，即

$$u_t \sim N(0,1)$$

式中，$u_t=\begin{cases} \varepsilon_t/\sqrt{h_t}, & GARCH \\ v_t/\sqrt{h_t}, & AR-GARCH \end{cases}$

1. 假设条件

H_0：u_t 服从标准正态分布$\leftrightarrow H_1$：u_t 不服从标准正态分布

2. 检验统计量

1982 年 Bera 和 Jarque 基于偏度和峰度，给出了正态检验统计量：

$$T_n=\frac{n}{6}b_1^2+\frac{n}{24}(b_2-3)^2$$

式中，b_1 为偏度函数；b_2 为峰度函数。

$$b_1 = \frac{\sqrt{n}\sum_{t=1}^{n}\hat{\mu}_t^3}{(\sum_{t=1}^{n}\hat{\mu}_t^2)^{\frac{3}{2}}}, b_2 = \frac{\sqrt{n}\sum_{t=1}^{n}\hat{\mu}_t^4}{(\sum_{t=1}^{n}\hat{\mu}_t^2)^2}$$

在原假设成立的条件下，$T_n \sim \chi^2(2)$。

3. 检验结果

检验结果如表 5—17 所示。

表 5—17

显著性水平	临界值	T_n 统计量	P 值	检验结果
α	$\chi^2_{1-\alpha}(2)$	$T_n \geqslant \chi^2_{1-\alpha}(2)$	$\leqslant \alpha$	拒绝 H_0
		$T_n < \chi^2_{1-\alpha}(1)$	$> \alpha$	接受 H_0

【例 5.12 续】 $AR(2)-GARCH(1,1)$模型正态检验结果如下：

$$T_n = 1.158\ 5，P = 0.560\ 3$$

检验结果显示接受 $v_t/\sqrt{h_t} \sim N$（0，1）假定。换言之，$AR(2)-GARCH(1,1)$ 模型拟合得比较充分，通过了模型检验。

查看拟合图，直观考察 $AR(2)-GARCH(1,1)$模型拟合效果，如图 5—44 所示。

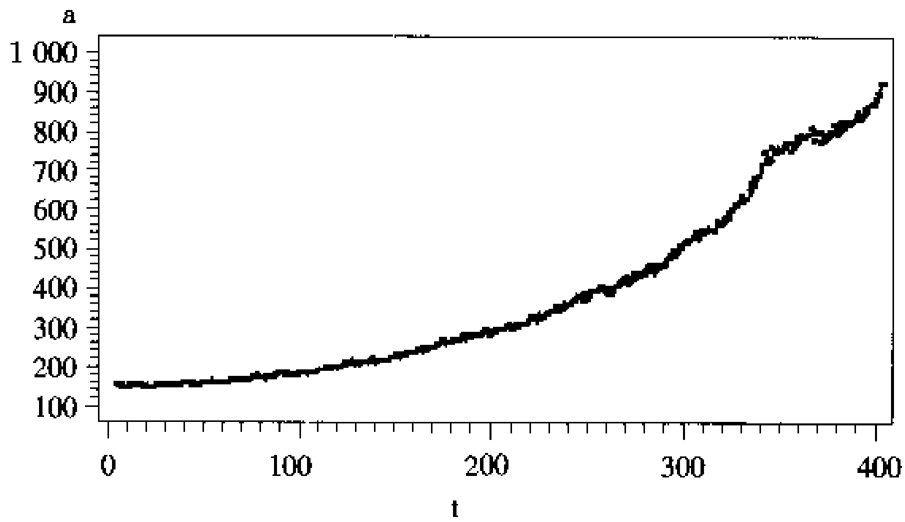


图 5—44 拟合效果图

图中，圆点为观察值序列，曲线为拟合值序列。拟合图显示点线几乎完全重合，这说明 $AR(2)-GARCH(1,1)$ 模型拟合得非常成功。

5.7 习　题

1. 某股票连续若干天的收盘价如表 5—18 所示。

表 5—18

304	303	307	299	296	293	301	293	301	295	284	286	286	287	284
282	278	281	278	277	279	278	270	268	272	273	279	279	280	275
271	277	278	279	283	284	282	283	279	280	280	279	278	283	278
270	275	273	273	272	275	273	273	272	273	272	273	271	272	271
273	277	274	274	272	280	282	292	295	295	294	290	291	288	288
290	293	288	289	291	293	293	290	288	287	289	292	288	288	285
282	286	286	287	284	283	286	282	287	286	287	292	292	294	291
288	289													

选择适当模型拟合该序列的发展，并估计下一天的收盘价。

2. 某城市连续 14 年的月度婴儿出生率数据如表 5—19 所示。

表 5—19

26.663	23.598	26.931	24.740	25.806	24.364	24.477	23.901
23.175	23.227	21.672	21.870	21.439	21.089	23.709	21.669
21.752	20.761	23.479	23.824	23.105	23.110	21.759	22.073
21.937	20.035	23.590	21.672	22.222	22.123	23.950	23.504
22.238	23.142	21.059	21.573	21.548	20.000	22.424	20.615
21.761	22.874	24.104	23.748	23.262	22.907	21.519	22.025
22.604	20.894	24.677	23.673	25.320	23.583	24.671	24.454
24.122	24.252	22.084	22.991	23.287	23.049	25.076	24.037
24.430	24.667	26.451	25.618	25.014	25.110	22.964	23.981
23.798	22.270	24.775	22.646	23.988	24.737	26.276	25.816
25.210	25.199	23.162	24.707	24.364	22.644	25.565	24.062
25.431	24.635	27.009	26.606	26.268	26.462	25.246	25.180
24.657	23.304	26.982	26.199	27.210	26.122	26.706	26.878
26.152	26.379	24.712	25.688	24.990	24.239	26.721	23.475
24.767	26.219	28.361	28.599	27.914	27.784	25.693	26.881
26.217	24.218	27.914	26.975	28.527	27.139	28.982	28.169
28.056	29.136	26.291	26.987	26.589	24.848	27.543	26.896
28.878	27.390	28.065	28.141	29.048	28.484	26.634	27.735
27.132	24.924	28.963	26.589	27.931	28.009	29.229	28.759
28.405	27.945	25.912	26.619	26.076	25.286	27.660	25.951
26.398	25.565	28.865	30.000	29.261	29.012	26.992	27.897

(1) 选择适当的模型拟合该序列的发展。

(2) 使用拟合模型预测下一年度该城市月度婴儿出生率。

3. 获得 100 个 $ARIMA(0,1,1)$序列观察值 $x_1,x_2,\cdots,x_{100}$。

(1) 已知 $\theta_1=0.3$，$x_{100}=50$，$\hat{x}_{100}(1)=51$，求 $\hat{x}_{100}(2)$ 的值。

(2) 假定新获得 $x_{101}=52$，求 $\hat{x}_{101}(1)$ 的值。

4. 1867—1938 年英国（英格兰及威尔士）绵羊数量如表 5—20 所示。

表 5—20

220 3	236 0	225 4	216 5	202 4	207 8	221 4	229 2	220 7	211 9	211 9	213 7
213 2	195 5	178 5	174 7	181 8	190 9	195 8	189 2	191 9	185 3	186 8	199 1
211 1	211 9	199 1	185 9	185 6	192 4	189 2	191 6	196 8	192 8	189 8	185 0
184 1	182 4	182 3	184 3	188 0	196 8	202 9	199 6	193 3	180 5	171 3	172 6
175 2	179 5	171 7	164 8	151 2	133 8	138 3	134 4	138 4	148 4	159 7	168 6
170 7	164 0	161 1	163 2	177 5	185 0	180 9	165 3	164 8	166 5	162 7	179 1

(1) 确定该序列的平稳性。

(2) 选择适当模型，拟合该序列的发展。

(3) 利用拟合模型预测 1939—1945 年英国绵羊的数量。

5. 1969 年 1 月—1994 年 9 月澳大利亚储备银行（Reserve Bank of Australia）2 年期有价证券月度利率数据如表 5—21 所示。

表 5—21

4.99	5.00	5.03	5.03	5.25	5.26	5.30	5.45	5.49	5.52	5.70
5.68	5.65	5.80	6.50	6.45	6.48	6.45	6.35	6.40	6.43	6.43
6.44	6.45	6.48	6.40	6.35	6.40	6.30	6.32	6.35	6.13	5.70
5.58	5.18	5.18	5.17	5.15	5.21	5.23	5.05	4.65	4.65	4.60
4.67	4.69	4.68	4.62	4.63	4.90	5.44	5.56	6.04	6.06	6.06
8.07	8.07	8.10	8.05	8.06	8.07	8.06	8.11	8.60	10.80	11.00
11.00	11.00	9.48	9.18	8.62	8.3	8.47	8.44	8.44	8.46	8.49
8.54	8.54	8.50	8.44	8.49	8.40	8.46	8.50	8.50	8.47	8.47
8.47	8.48	8.48	8.54	8.56	8.39	8.89	9.91	9.89	9.91	9.91
9.90	9.88	9.86	9.86	9.74	9.42	9.27	9.26	8.99	8.83	8.83
8.83	8.82	8.83	8.83	8.79	8.79	8.69	8.66	8.67	8.72	8.77
9.00	9.61	9.70	9.94	9.94	9.94	9.95	9.94	9.96	9.97	10.83
10.75	11.20	11.40	11.54	11.50	11.34	11.50	11.50	11.58	12.42	12.85
13.10	13.12	13.10	13.15	13.10	13.20	14.20	14.75	14.60	14.60	14.45
14.50	14.80	15.85	16.20	16.50	16.40	16.40	16.35	16.10	13.70	13.50
14.00	12.30	12.00	14.35	14.60	12.50	12.75	13.70	13.45	13.55	12.60
12.00	11.00	11.60	12.05	12.35	12.70	12.45	12.55	12.20	12.10	11.15
11.85	12.10	12.50	12.90	12.50	13.20	13.65	13.65	13.50	13.45	13.35
14.45	14.30	15.05	15.55	15.65	14.65	14.15	13.30	12.65	12.70	12.80

续前表

14.50	15.10	15.15	14.30	14.25	14.05	14.70	15.05	14.05	13.80	13.25
13.00	12.85	12.60	11.80	13.00	12.35	11.45	11.35	11.55	10.85	10.90
12.30	11.70	12.05	12.30	12.90	13.05	13.30	13.85	14.65	15.05	15.15
14.85	15.70	15.40	15.10	14.80	15.80	15.80	15.00	14.40	13.80	14.30
14.15	14.45	14.10	14.05	13.75	13.30	13.00	12.55	12.25	11.85	11.50
11.10	11.15	10.70	10.25	10.55	10.25	10.30	9.60	8.40	8.20	7.25
8.35	8.25	8.30	7.40	7.15	6.35	5.65	7.40	7.20	7.05	7.10
6.85	6.50	6.25	5.95	5.65	5.85	5.45	5.30	5.20	5.55	5.15
5.40	5.35	5.10	5.80	6.35	6.50	6.95	8.05	7.85	7.75	8.60

(1) 考察该序列的方差齐性。

(2) 选择适当的模型拟合该序列的发展。

5.8 上机指导

5.8.1 拟合 *ARIMA* 模型

由于 *ARMA* 模型是 *ARIMA* 模型的一种特例，所以在 *SAS* 系统中这两种模型的拟合都放在了 *ARIMA* 过程中。我们已经在第 3 章进行 *ARMA* 模型拟合时介绍了 *ARIMA* 过程的基本命令格式。在此以临时数据集 example5_1 的数据为例介绍 *ARIMA* 模型拟合与 *ARMA* 模型拟合的不同之处。

```
data example5_1;
input x@@;
difx=dif (x);
t=_n_;
cards;
 1.05   -0.84   -1.42    0.20    2.81    6.72    5.40    4.38
 5.52    4.46    2.89   -0.43   -4.86   -8.54  -11.54  -16.22
-19.41  -21.61  -22.51  -23.51  -24.49  -25.54  -24.06  -23.44
-23.41  -24.17  -21.58  -19.00  -14.14  -12.69   -9.48  -10.29
 -9.88   -8.33   -4.67   -2.97   -2.91   -1.86   -1.91   -0.80
proc gplot;
plot x * t;
symbol v=star c=black i=join;
run;
```

输出时序图显示这是一个典型的非平稳序列。如图 5—45 所示。

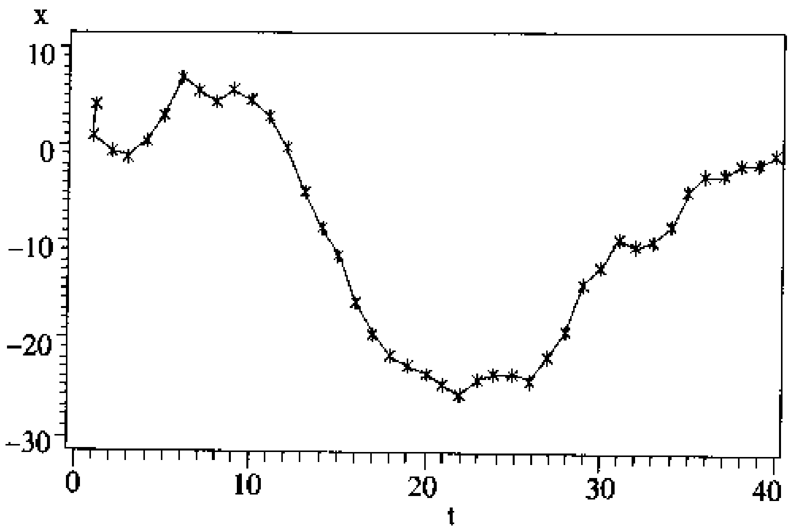


图 5—45 序列 x 时序图

考虑对该序列进行 1 阶差分运算，同时考察差分后序列的平稳性，在原程序基础上添加相关命令，程序修改如下：

```
data example5_1;
input x@@;
difx=dif (x);
t=_n_;
cards;
1.05   -0.84   -1.42   0.20   2.81   6.72   5.40   4.38
5.52   4.46   2.89   -0.43   -4.86   -8.54   -11.54   -16.22
-19.41   -21.61   -22.51   -23.51   -24.49   -25.54   -24.06   -23.44
-23.41   -24.17   -21.58   -19.00   -14.14   -12.69   -9.48   -10.29
-9.88   -8.33   -4.67   -2.97   -2.91   -1.86   -1.91   -0.80
proc gplot;
plot x*t difx*t;
symbol v=star c=black i=join;
proc arima;
identify var=x(1);
estimate p=1;
forecast lead=5 id=t;
run;
```

语句说明：

(1) DATA 步中的命令"difx=dif(x);"，这是指令系统对变量 x 进行 1 阶差分，差分后的序列值赋值给变量 difx。其中 dif () 是差分函数，假如要差分的变量名为 x，常见的几种差分表示为：

1 阶差分：dif(x)

2 阶差分：dif(dif(x))

k 步差分：difk(x)

(2) 我们在 GPLOT 过程中添加绘制了一个时序图"difx * t"，这是为了直观考察 1 阶差分后序列的平稳性。所得时序图如图 5—46 所示。

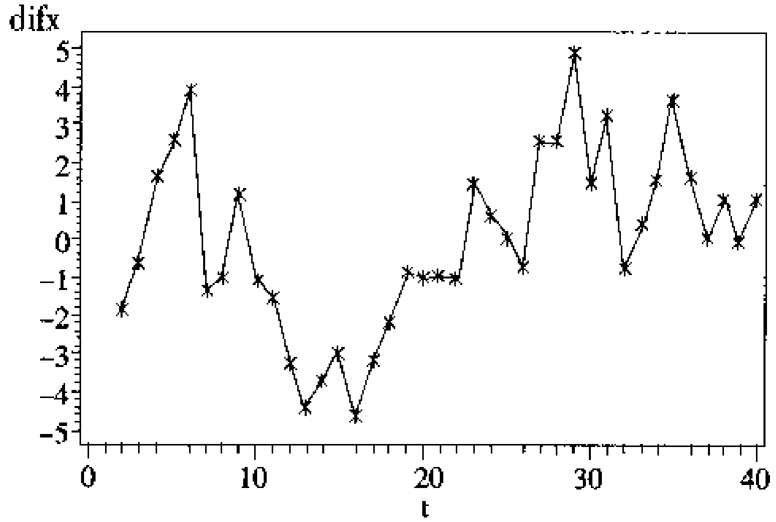


图 5—46　序列 difx 时序图

时序图显示差分后序列 difx 没有明显的非平稳特征。

(3) "identify var=x(1);"，使用该命令可以识别差分后序列的平稳性、纯随机性和适当的拟合模型阶数。其中 x(1)表示识别变量 x 的 1 阶差分后序列。SAS 支持多种形式的差分序列识别：

var=x(1)，表示识别变量 x 的 1 阶差分后序列 ∇x_t；

var=x(1,1)，表示识别变量 x 的 2 阶差分后序列 $\nabla^2 x_t$；

var=x(k)，表示识别变量 x 的 k 步差分后序列 $\nabla_k x_t$；

var=x(k,s)，表示识别变量 x 的 k 步差分后，再进行 s 步差分后序列 $\nabla_k \nabla_s x_t$。

识别部分的输出结果显示 1 阶差分后序列 difx 为平稳非白噪声序列，且具有显著的自相关系数不截尾，偏自相关系数 1 阶截尾的性质。

(4) "estimate p=1;" 对 1 阶差分后序列 ∇x_t 拟合 $AR(1)$ 模型。输出拟合结果显示常数项不显著，添加或修改估计命令如下：

estimate p=1 noint;

这是命令系统不要常数项拟合 AR(1) 模型，拟合结果显示模型显著且参数显著。如图 5—47 所示。

Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr> |t|	Lag
AR1, 1	0.669 33	0.12 118	5.52	<.000 1	1

Variance Estimate	2.987 777
Std Error Estimate	1.728 519
AIC	154.350 8
SBC	156.014 4
Number of Residuals	

* AIC and SBC do not include log determinant.

Autocorrelation Check of Residuals

To Lag	ChiSquare	DF	Pr> ChiSq	Autocorrelations					
6	5.60	5	0.339 0	−0.087	0.120	−0.160	0.171	0.109	0.188
12	7.48	11	0.729 3	−0.163	0.001	−0.021	−0.065	−0.049	−0.038
18	10.31	17	0.890 2	−0.106	0.043	−0.104	−0.118	−0.067	0.012
24	12.85	23	0.955 2	−0.026	−0.023	−0.098	−0.073	0.096	−0.039

图 5—47　序列 difx 模型拟合结果

输出结果显示序列 x_t 的拟合模型为 *ARIMA* (1, 1, 0) 模型，模型口径为：

$$\nabla x_t = \frac{\varepsilon_t}{1-0.669\,33B}$$

或等价记为：

$$x_t = 1.669\,33x_{t-1} - 0.669\,33x_{t-2} + \varepsilon_t$$

(5) “forecast lead=5 id=t;”，利用拟合模型对序列 x_t 作 5 期预测。

5.8.2　拟合 *Auto-Regressive* 模型

在 SAS 系统中有一个 AUTOREG 程序，可以进行残差自回归模型拟合。下面以临时数据集 example5_2 的数据为例，介绍相关命令的使用。

一、建立数据集，绘制时序图

```
data example5_2;
input x@@;
t=_n_;
cards;
```

3.03	8.46	10.22	9.80	11.96	2.83
8.43	13.77	16.18	16.84	19.57	13.26
14.78	24.48	28.16	28.27	32.62	18.44
25.25	38.36	43.70	44.46	50.66	33.01
39.97	60.17	68.12	68.84	78.15	49.84
62.23	91.49	103.20	104.53	118.18	77.88
94.75	138.36	155.68	157.46	177.69	117.15

```
;
proc gplot data=example5 _ 2;
plot x * t=1;
symbol1 c=black i=join v=star;
run;
```

输出时序图如图 5—48 所示。

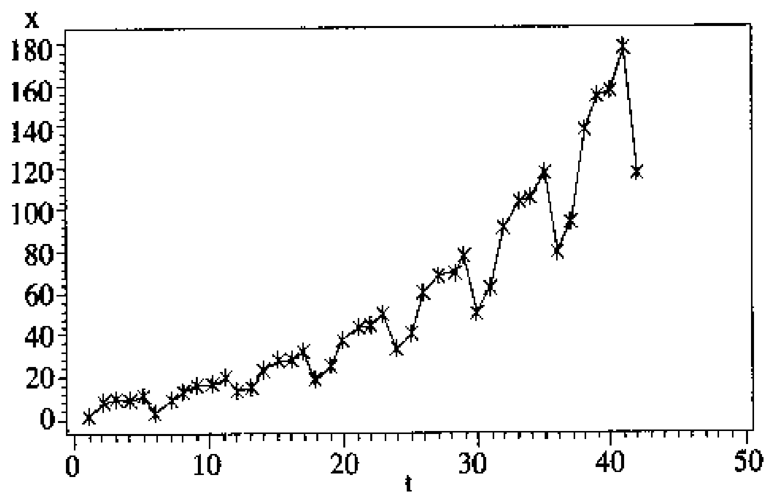


图 5—48 序列 x 的时序图

时序图显示序列 x 有一个明显的随时间线性递增的趋势，同时又有一定规律性的波动，所以不妨考虑使用误差自回归模型拟合该序列的发展。

二、因变量关于时间的回归模型

```
proc autoreg data=example5 _ 2;
model x=t/ dwprob;
run;
```

语句说明：

(1) "proc autoreg data=example5 _ 2;" 指令 SAS 系统对临时数据集 example5 _ 2 进行自回归程序分析。

(2) "model x=t/ dwprob;" 指令 SAS 系统以变量 t 作为自变量，变量 x 作

为因变量，建立线性模型：

$$x_t = a + bt + u_t$$

并给出残差序列 $\{u_t\}$ DW 检验统计量的分位点。

本例中，序列 x 关于变量 t 的线性回归模型最小二乘估计输出结果如图 5—49 所示。

Ordinary Least Squares Estimates			
SSE	15 464.524 8	DFE	40
MSE	386.613 12	Root MSE	19.662 48
SBC	374.828 818	AIC	371.353 479
Regress R-Square	0.830 0	Total R-Square	0.830 0
Durbin-Watson	0.762 8	Pr <DW	<.000 1
Pr> DW	1.000 0		

图 5—49　序列 x 关于变量 t 的线性回归模型最小二乘估计结果

本例输出结果显示 DW 统计量的值等于 0.762 8，输出概率显示残差序列显著正相关。所以应该考虑对残差序列拟合自相关模型，修改 AUTOREG 程序如下：

```
proc autoreg data=example5_2;
model x=t/ nlag=5 backstep method=ml;
run;
```

model 语句是指令系统对线性回归模型 $x_t = a + bt + u_t$ 的残差序列 $\{u_t\}$ 显示延迟 5 阶的自相关图，并拟合延迟 5 阶自相关模型，特别注意，SAS 输出的自回归模型结构为：

$$u_t = -\phi_1 u_{t-1} - \cdots - \phi_5 u_{t-5} + \varepsilon_t$$

即输出的自回归参数值与我们习惯定义的自回归参数值相差一个负号。

由于自相关延迟阶数的确定是由我们尝试选择的，所以 nlag 的阶数通常会指定得大一些。这就导致残差自回归模型中可能有部分参数不显著，因而添加逐步回归选项 backstep，指令系统使用逐步回归的方法筛选出显著的自相关因子，并使用极大似然的方法进行参数估计。

输出如下四方面的结果：

1. 因变量说明

因变量说明如图 5—50 所示。

The AUTOREG Procedure
Dependent Variable x

图 5—50　因变量说明

2. 普通最小二乘估计结果

该部分输出信息包括误差平方和（SSE）、自由度（DFE）、均方误（MSE）、根号均方误（Root MSE）、SBC 信息量、AIC 信息量、回归部分相关系数平方（Regress R-Square）、总的相关系数平方（Totel R Square），DW 统计量（Durbin-Watson）及所有待估参数的自由度、估计值、标准差、t-值和 t 统计量的 P 值。如图 5--51 所示。

Ordinary Least Squares Estimates

SSE	15 464.524 8	DFE	40
MSE	386.613 12	Root MSE	19.662 48
SBC	374.828 818	AIC	371.353 479
Regress R-Square	0.830 0	Total R-Square	0.830 0
Durbin-Watson	0.762 8		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr> |t|
Intercept	1	−20.910 2	6.178 0	−3.38	0.001 6
t	1	3.497 7	0.250 3	13.97	<.000 1

图 5—51　普通最小二乘估计结果

3. 回归误差分析

该部分共输出四方面的信息：残差序列自相关图、逐步回归消除的不显著项报告、初步均方误差（MSE）、自回归参数估计值。

本例该部分输出结果如图 5—52 所示。

Estimates of Autocorrelations

Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
0	368.2	1.000 000	|******************** |
1	221.9	0.602 573	|************ |
2	116.1	0.315 203	|****** |
3	57.630 1	0.156 517	|*** |
4	21.814 5	0.059 246	|* |
5	74.864 4	0.203 324	|**** |

Backward Elimination of Autoregressive Terms

Lag	Estimate	t Value	Pr>|t|
3	-0.035 768	-0.19	0.853 7
2	0.061 250	0.37	0.713 2
4	0.216 740	1.37	0.180 1
5	-0.168 214	-1.33	0.192 5

Preliminary MSE　　234.5

Estimates of Autoregressive Parameters

Lag	Coefficient	Standard Error	t Value
1	-0.602 573	0.127 793	-4.72

图 5—52　自回归误差分析输出结果

本例输出的残差序列自相关图显示残差序列有非常显著的1阶正相关。逐步回归向后消除报告显示除了延迟1阶的序列值显著自相关外，延迟其他阶数的序列值均不具有显著的自相关性，因此延迟2阶～5阶的自相关项被消除。初步均方误差为234.5，1阶残差自回归模型的参数为 $\hat{\phi}_1=-0.602\ 573$。特别注意，这意味着输出的自回归模型结果为：

$$u_t=0.602\ 573u_{t-1}+\varepsilon_t$$

4. 最终拟合模型

该部分输出三方面的汇总信息：收敛状况、极大似然估计结果和回归系数估计。本例该部分输出结果如图5—53所示。

Algorithm converged.

Maximum Likelihood Estimates

SSE	10 037.807	DFE	40
MSE	250.945 18	Root MSE	15.841 25
SBC	357.318 944	AIC	353.843 605
Regress R-Square	0.687 6	Total R-Square	0.953 3
Durbin-Watson	1.697 5		

NOTE: No intercept term is used. R-squares are redefined.

Variable	DF	Estimate	Standard Error	t Value	Approx Pr> |t|
t	1	2.763 8	0.294 8	9.38	<.000 1
AR1	1	−0.688 3	0.112 4	−6.12	<.000 1

Autoregressive parameters assumed given.

Variable	DF	Estimate	Standard Error	t Value	Approx Pr> |t|
t	1	2.763 8	0.294 6	9.38	<.000 1

图5—53 最终拟合模型输出结果

本例得到最终拟合模型为：

$$\begin{cases} x_t=2.763\ 8t+u_t \\ u_t=0.688\ 3u_{t-1}+\varepsilon_t,\quad \varepsilon_t \overset{i.i.d}{\sim} N(0,\ 250.945\ 18) \end{cases}$$

为了得到直观的拟合效果，还可以利用OUTPUT命令将拟合结果存入SAS数据集中，并对输出结果作图，相关命令如下：

```
proc autoreg data=example5_2;
model x=t/nlag=5 backstep method=ml noint;
output out=out p=xp pm=trend;
```

```
proc gplot data=out;
plot x * t=2 xp * t=3 trend * t=4 / overlay;
symbol2 v=star i=nnone c=black;
symbol3 v=none i=join c=red w=2 l=3;
symbol4 v=none i=join c=green w=2;
run;
```

语句说明：

“output out=out p=xp pm=trend;”，该命令是指令系统将部分结果输入临时数据集 OUT，选择输出的第一个信息为整体模型的拟合值（P 选项），该拟合变量取名为 XP；选择输出的第二个信息为线性趋势拟合值（PM 选项），还可以选择 R 选项输出拟合残差项，本例不要求输出此项。

输出图像如图 5—54 所示。

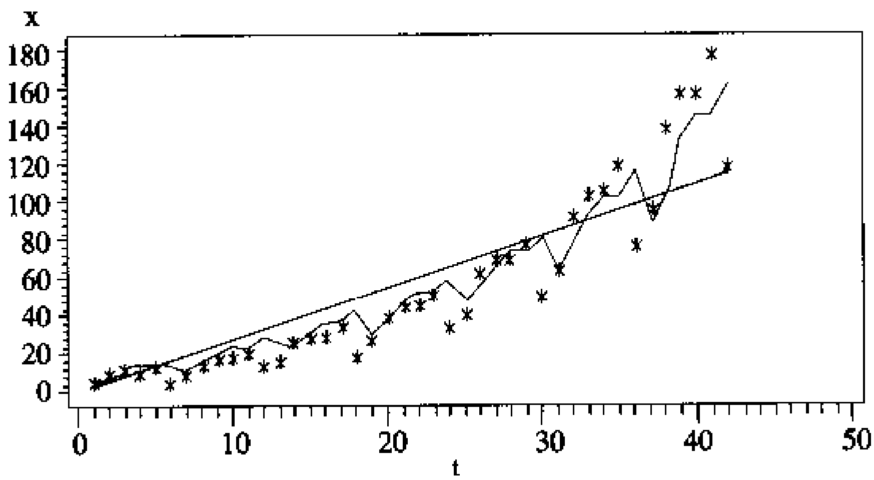


图 5—54　拟合效果图

三、延迟因变量回归模型

```
proc autoreg data=example5_2;
model x=lagx/lagdep=lagx;
run;
```

语句说明：

(1) 首先在 DATA 步中添加命令“lagx=lag(x);”，该语句指令系统使用延迟函数生成序列 x 的 1 阶延迟序列，并将该序列赋值给变量 lagx，即

$$\text{lag}x_t=x_{t-1}，\ t\geqslant 2$$

(2) “model x=lagx/lagdep=lagx;”指令系统建立带有延迟因变量的回归模型

$x_t = a + bx_{t-1} + u_t$

并通过 LAGDEP 选项指定被延迟的因变量名。本例输出结果如图 5—55 所示。

Dependent Variable x

Ordinary Least Squares Estimates

SSE		11 462.72	DFE		39
MSE		293.915 90	Root MSE		17.143 98
SBC		354.744 717	AIC		351.317 573
Regress R-Square		0.870 1	Total R-Square		0.870 1
Durbin h		−0.291 6	Pr ＜h		0.385 3
			Standard		Approx
Variable	DF	Estimate	Error	t Value	Pr＞ |t|
Intercept	1	5.945 9	4.072 1	1.46	0.152 3
Lagx	1	0.940 1	0.058 2	16.16	＜.000 1

图 5—55　带延迟因变量回归分析结果

由于带有延迟因变量，所以这种场合在回归模型估计结果中输出的是 Durbin h 统计量。本例中 Durbin h 统计量的分布函数达到0.385 3，这表示残差序列不存在显著的相关性。不需要考虑对残差序列继续拟合自回归模型。

再注意参数检验结果，如图 5—56 所示。

Variable	DF	Estimate	Standard Error	t Value	Approx Pr＞ |t|
Intercept	1	5.945 9	4.072 1	1.46	0.152 3
Lagx	1	0.940 1	0.058 2	16.16	＜.000 1

图 5—56　参数估计结果

在显著性水平默认为 0.05 的条件下，截距项不显著（P 值大于 0.152 3），所以可以考虑在模型拟合命令中增加 NOINT 选项。最后输出拟合结果，并绘制拟合时序图，相关命令如下：

```
proc autoreg data=example5_2;
model x=lagx/lagdep=lagx noint;
output out=out p=xp;
proc gplot data=out;
plot x*t=2 xp*t=3 / overlay;
symbol2 v=star i=none c=black;
symbol3 v=none i=join c=red w=2;
run;
```

模型拟合部分输出的信息表示最终的拟合模型为：

$$x_t = 1.004x_{t-1} + \varepsilon_t, \varepsilon_t \overset{i.i.d}{\sim} N(0,302.234\,37)$$

输出的拟合效果图如图 5—57 所示。

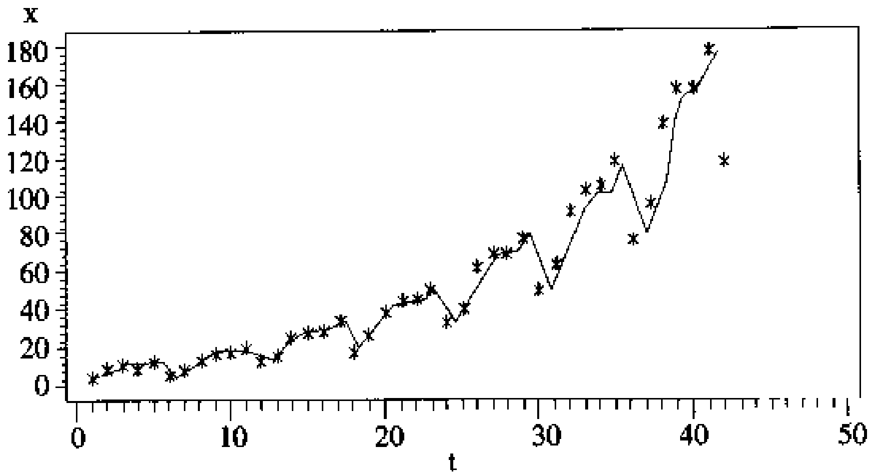


图 5—57　带有延迟因变量的回归模型拟合效果图

本例中由于没有残差自回归项，最终拟合值就是趋势值，所以只需绘制其中之一即可。

5.8.3　拟合 *GARCH* 模型

SAS 系统中 AUTOREG 过程功能非常强大，不仅可以提供如上的分析功能，还可以提供异方差性检验乃至于条件异方差模型建模。以临时数据集 example5 _ 3 数据为例，介绍 *GARCH* 模型的拟合，相关程序如下：

```
data example5 _ 3;
input x@@;
t= _ n _ ;
cards;
```

10.77	13.30	16.64	19.54	18.97	20.52	24.36
23.51	27.16	30.80	31.84	31.63	32.68	34.90
33.85	33.09	35.46	35.32	39.94	37.47	35.24
33.03	32.67	35.20	32.36	32.34	38.45	38.17
32.14	39.70	49.42	47.86	48.34	62.50	63.56
67.61	64.59	66.17	67.50	76.12	79.31	78.85
81.34	87.06	86.41	93.20	82.95	72.96	61.10

61.27	71.58	88.34	98.70	97.31	97.17	91.17
80.20	85.12	81.40	70.87	57.75	52.35	67.50
87.95	85.46	84.55	98.16	102.42	113.02	119.95
122.37	126.96	122.79	127.96	139.20	141.05	140.87
137.08	145.53	145.59	134.36	122.54	106.92	97.23
110.39	132.40	152.30	154.91	152.69	162.67	160.31
142.57	146.54	153.83	141.81	157.83	161.79	142.07
139.43	140.92	154.61	172.33	191.78	199.27	197.57
189.29	181.49	166.84	154.28	150.12	165.17	170.32

```
;
proc gplot data=example5_3;
plot x * t=1;
symbol1 c=black i=join v=star;
proc autoreg data=example5_3;
model x=t/nlag=5 dwprob archtest;
model x=t/nlag=2 noint garch= (p=1, q=1);
output out=out p=xp;
proc gplot data=out;
plot x * t=2 xp * t=3/overlay;
symbol2 v=star i=none c=black;
symbol3 v=none i=join c=red w=2;
run;
```

该序列输出时序图如图 5—58 所示。

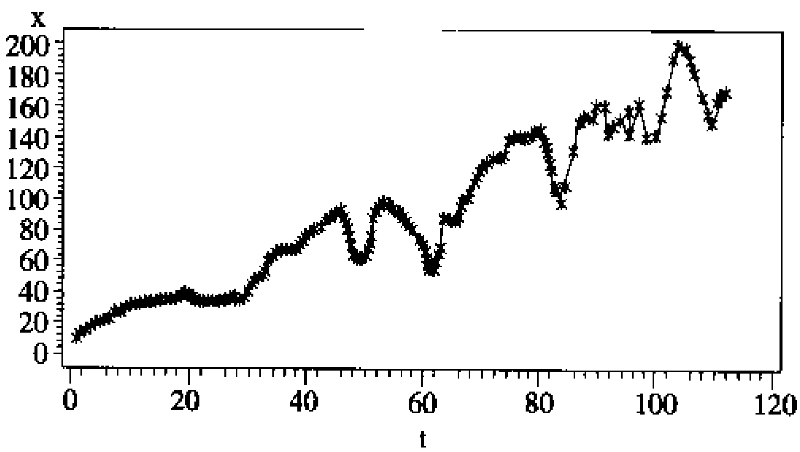


图 5—58 序列时序图

时序图显示序列具有显著线性递增趋势，且波动幅度随时间递增，所以考虑使用 AUTOREG 过程建立序列 $\{x_t\}$ 关于时间 t 的线性回归模型，并检验残差序列的自相关性和异方差性，如果检验结果显示残差序列具有显著自相关性，建立残差自回归模型；如果残差序列具有显著的异方差性，则要建立条件异方差模型。

语句说明：

(1) “model x=t/nlag=5 dwprob archtest;”，该命令指令系统建立序列 $\{x_t\}$ 关于时间 t 的线性回归模型，并检验残差序列 5 阶延迟的自相关性并输出 DW 检验的 P 值，同时对残差序列进行异方差检验。

DW 检验结果显示残差序列具有显著的正自相关性，如图 5—59 所示。

Ordinary Least Squares Estimates			
SSE	24 013.858 6	DFE	110
MSE	218.307 81	Root MSE	14.775 24
SBC	928.482 631	AIC	923.045 633
Regress R-Square	0.918 4	Total R-Square	0.918 4
Durbin-Watson	0.328 0	Pr <DW	<.000 1
Pr> DW	1.000 0		

图 5—59　普通最小二乘估计输出结果

残差序列 5 阶延迟自相关图显示残差序列至少具有 2 阶显著自相关性，如图 5—60 所示。

```
                         Estimates of Autocorrelations
Lag  Covariance  Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
 0      214.4     1.000 000 |                    |********************|
 1      179.1     0.835 094 |                    |*****************   |
 2      116.9     0.545 110 |                    |***********         |
 3    58.468 3    0.272 694 |                    |*****               |
 4     7.794 3    0.036 352 |                    |*                   |
 5   -22.908 0   -0.106 843 |                  **|                    |
```

图 5—60　残差序列自相关图

参数估计结果显示回归模型常数截距项不显著，如图 5—61 所示。

Variable	DF	Estimate	Standard Error	t Value	Approx Pr> |t|
Intercept	1	5.928 0	5.299 6	1.12	0.265 9
t	1	1.521 2	0.081 0	18.78	<.000 1

图 5—61　线性回归模型参数估计结果

异方差检验结果显示残差序列具有显著的异方差性，且具有显著的长期相关性，如图 5—62 所示。

Q and LM Tests for ARCH Disturbances				
Order	Q	Pr> Q	LM	Pr> LM
1	56.897 2	<.000 1	55.844 2	<.000 1
2	64.843 8	<.000 1	67.844 4	<.000 1
3	66.018 0	<.000 1	74.798 2	<.000 1
4	67.129 7	<.000 1	76.609 4	<.000 1
5	67.504 7	<.000 1	76.741 1	<.000 1
6	67.546 5	<.000 1	76.850 8	<.000 1
7	68.250 9	<.000 1	76.936 5	<.000 1
8	69.558 3	<.000 1	76.938 4	<.000 1
9	70.665 5	<.000 1	76.946 4	<.000 1
10	71.557 7	<.000 1	77.113 4	<.000 1
11	71.919 6	<.000 1	77.840 2	<.000 1
12	71.924 9	<.000 1	77.851 2	<.000 1

图 5—62 普通最小二乘估计输出结果

(2) "model x=t/nlag=2 noint garch= (p=1, q=1);"，综合考虑残差序列自相关性和异方差性检验结果，尝试拟合无回归常数项的 *AR*(2) − *GARCH*(1, 1) 模型。

模型最终拟合结果如图 5--63 所示。

The AUTOREG Procedure

GARCH Estimates

SSE	5 759.617 13	Observations	112
MSE	51.425 15	Uncond Var	216.158 433
Log Likelihood	−368.380 39	Total R-Square	0.995 4
SBC	765.071 78	AIC	748.760 787
Normality Test	2.3384	Pr> ChiSq	0.310 6

NOTE: No intercept term is used. R-squares are redefined.

Variable	DF	Estimate	Standard Error	t Value	Approx Pr> |t|
t	1	1.644 3	0.065 7	25.03	<.000 1
AR1	1	−1.283 9	0.098 8	−12.99	<.000 1
AR2	1	0.429 2	0.095 7	4.48	<.000 1
ARCH0	1	1.593 4	2.038 7	0.78	0.434 5
ARCH1	1	0.281 7	0.159 6	1.76	0.077 6
GARCH1	1	0.7140 9	0.138 0	5.15	<.000 1

图 5—63 普通最小二乘估计输出结果

参数检验结果显示除 $GARCH(1, 1)$ 模型中的常数项不显著外，其他变量均显著，整个模型的 R^2 高达 0.995 4，且正态性检验不显著（P 值为 0.310 6），这与假定 $GARCH$ 的残差函数 $\varepsilon_t/\sqrt{h_t}$ 服从正态分布相吻合，所以可以认为该模型拟合成功。

最终模型口径为：

$$\begin{cases} x_t = 1.6443t + u_t \\ u_t = 1.2839u_{t-1} - 0.4292u_{t-2} + \varepsilon_t \\ \varepsilon_t = \sqrt{h_t}a_t, a_t \overset{i.i.d}{\sim} N(0, 51.42515) \\ h_t = 1.5934 + 0.2817\varepsilon_{t-1}^2 + 0.7109h_{t-1} \end{cases}$$

最终输出拟合效果图如图 5- 64 所示。

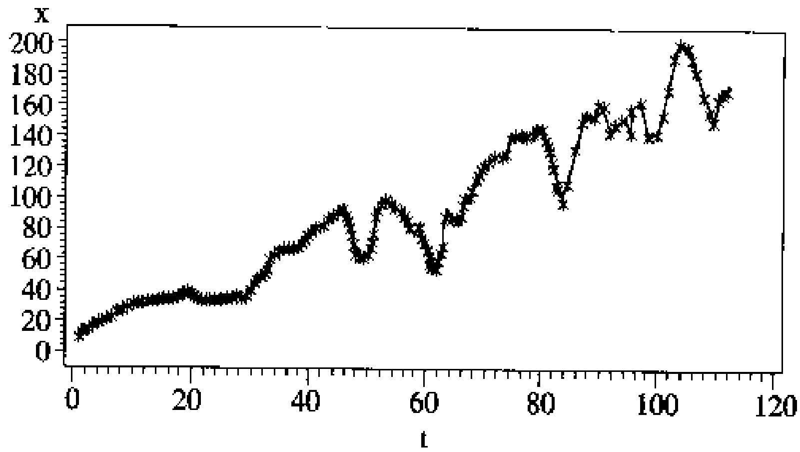


图 5—64 *GARCH* 模型拟合效果图

21世纪 教材

第6章

多元时间序列分析

在前面几章中我们介绍的都是一元时间序列的分析方法。实际上，很多序列的变化规律都会受到其他序列的影响。比如说当我们分析居民消费支出序列时，消费会受到收入的显著影响，如果我们能将收入也纳入研究范围，就能得到更精确的消费预测。这就涉及多元时间序列分析。

对多元时间序列的分析很早就开始了。1976年，Cox和Jenkins在*Time Series Analysis：Forecasting and Control*一书中就采用过将天然气的输入速率作为输入变量，研究CO_2的输出浓度，由此将时间序列的分析领域由一元拓展到了多元的场合。

但是从技术上来讲，当时要求输入序列和研究序列都是平稳的。显然响应序列和输入序列均平稳的要求是非常苛刻的，这严重限制了多元时间序列分析的运用和发展。直到1987年Engle和Granger提出了协整（cointegration）的概念。

在协整理论下，并不要求响应序列和输入序列自身平稳，只要求它们的回归残差序列平稳。残差序列平稳比响应序列与输入序列均平稳容易实现多了。这个概念的提出极大地促进了多元时间序列分析的发展。它实际上是将多元回归分析和时间序列分析有机地结合在了一起，有效地提高了预测的精度。

本章将介绍多元时间序列建模的原理和方法。

6.1　平稳多元序列建模

1976 年，Cox 和 Jenkins 采用带输入变量的 *ARIMA* 模型，为平稳多元序列建模。

该模型的构造思想是：假设响应序列 $\{y_t\}$ 和输入变量序列（即自变量序列）$\{x_{1t}\}$，$\{x_{2t}\}$，…，$\{x_{kt}\}$ 均平稳，首先构建响应序列和输入变量序列的回归模型：

$$y_t = \mu + \sum_{k=1}^{k} \frac{\Theta_i(B)}{\Phi_i(B)} B^{l_i} x_{it} + \varepsilon_t$$

式中，$\Phi_i(B)$为第 i 个输入变量的自回归系数多项式；$\Theta_i(B)$为第 i 个输入变量的移动平均系数多项式；l_i 为第 i 个输入变量的延迟阶数；$\{\varepsilon_t\}$ 为回归残差序列。

因为 $\{y_t\}$ 和 $\{x_{1t}\}$，$\{x_{2t}\}$，…，$\{x_{kt}\}$ 均平稳，平稳序列的线性组合仍然是平稳的，所以残差序列 $\{\varepsilon_t\}$ 为平稳序列：

$$\varepsilon_t = y_t - \left(\mu + \sum_{k=1}^{k} \frac{\Theta_i(B)}{\Phi_i(B)} B^{l_i} x^{it}\right)$$

使用 *ARMA* 模型继续提取残差序列 $\{\varepsilon_t\}$ 中的相关信息。最终得到的模型形为：

$$\begin{cases} y_t = \mu + \sum_{k=}^{k} \frac{\Theta_i(B)}{\Phi_i(B)} B^{l_i} x_{it} + \varepsilon_t \\ \varepsilon_t = \frac{\Theta(B)}{\Phi(B)} a_t \end{cases} \tag{6.1}$$

模型（6.1）被称为动态回归模型，简记为 *ARIMAX*。式中，$\Phi(B)$为残差序列自回归系数多项式；$\Theta(B)$为残差序列移动平均系数多项式；a_t 为零均值白噪声序列。

【例 6.1】 在天然气炉中，输入的是天然气，输出的是 CO_2，CO_2 的输出浓度与天然气的输入速率有关。现在以中心化后的天然气输入速率为输入序列，建立 CO_2 的输出百分浓度模型（数据见附录 1.20）。

时序图直观显示输入序列和输出序列均平稳，如图 6—1、图 6—2 所示。

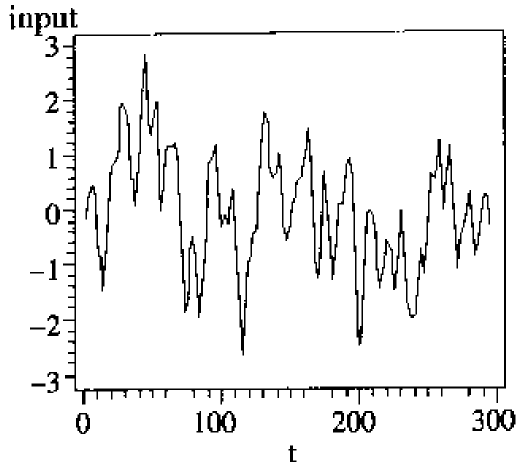


图 6—1　输入序列时序图

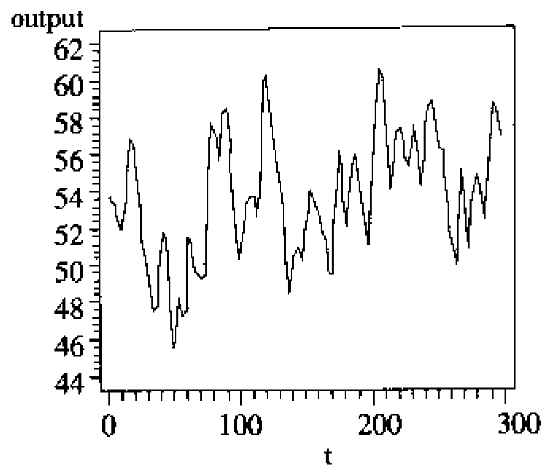


图 6—2　输出序列时序图

如果不考虑输入序列和输出序列之间的相关性，将它们作为两个一元时间序列分析，容易验证输入序列 $AR(3)$为模型：

$$x_t=-0.1228+\frac{a_t}{1-1.97607B+1.37499B^2-0.34336B^3}$$

输出序列为 $AR(1，2，4)$疏系数模型：

$$y_t=53.90176+\frac{a_t}{1-3.10703B+1.34005B^2-0.21274B^4}$$

且输出序列模型的 AIC=196.312 1，SBC=211.073 5。

考虑到输入天然气速率与输出 CO_2 的浓度之间有密切的联系，将输入天然气速率作为输入变量考虑进输出序列的模型中。通过协相关图，考察回归模型的结构。其中延迟 k 协方差函数定义为：

$$\mathrm{Cov}_k=\mathrm{Cov}(y_t,x_{t-k})=E[(y_t-Ey_t)(x_{t-k}-Ex_{t-k})]$$

延迟 k 协相关系数为：

$$C\rho_k=\frac{\mathrm{Cov}(y_t,x_{t-k})}{\mathrm{Var}(y_t)\mathrm{Var}(x_{t-k})}$$

本例中，考察协相关图（图 6—3）的性质可以看出延迟 3 阶～7 阶的协相关系数显著非零，这说明输出序列和输入序列之间有至少 3 期的滞后效应，且回归模型可以用如下模型表示：

$$y_t=\mu+\omega_3x_{t-3}+\omega_4x_{t-4}+\omega_5x_{t-5}+\omega_6x_{t-6}+\omega_7x_{t-7}+\varepsilon_t$$

当显著延迟阶数比较多时，可以采用 $ARMA$ 模型结构，以减少待估参数的个数。如上回归模型不妨采用如下 $ARMA$（1，2）结构表达：

$$y_1=\mu+\frac{\theta_0-\theta_1B-\theta_2B^2}{1-\phi_1B}B^3x_t+\varepsilon_t$$

```
                              Crosscorrelations
Lag  Covariance      Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
-10   0.001 568 3    0.023 01   |                   .  |  .                   |
 -9   0.000 135 02   0.001 98   |                   .  |  .                   |
 -8  -0.006 048 0   -.088 72    |                   ** |  .                   |
 -7  -0.001 762 4   -.025 85    |                   .* |  .                   |
 -6  -0.008 053 9   -.118 14    |                   ** |  .                   |
 -5  -0.000 094 4   -.001 38    |                   .  |  .                   |
 -4  -0.001 280 2   -.018 78    |                   .  |  .                   |
 -3  -0.003 107 8   -0.455 9    |                   .* |  .                   |
 -2   0.000 652 12   0.009 57   |                   .  |  .                   |
 -1  -0.001 916 6   -.028 11    |                   .* |  .                   |
  0  -0.000 367 3   -.005 39    |                   .  |  .                   |
  1   0.003 893 9    0.057 12   |                   .  |* .                   |
  2  -0.001 697 1   -.024 89    |                   .  |  .                   |
  3    -0.019 231   -.282 10    |               ****** |  .                   |
  4    -0.022 479   -.329 74    |              ******* |  .                   |
  5    -0.030 909   -.453 41    |            ********* |  .                   |
  6    -0.018 122   -.265 83    |                ***** |  .                   |
  7    -0.011 426   -.167 61    |                  *** |  .                   |
  8  -0.001 735 5   -.025 46    |                   .* |  .                   |
  9   0.002 259 0    0.033 14   |                   .  |* .                   |
 10  -0.003 515 2   -.051 56    |                   .* |  .                   |
                       "."marks two standard errors
```

图 6—3 输出序列与输入序列的协相关图

使用观察值序列确定回归模型的口径为：

$$y_t=53.322\ 56+\frac{-0.564\ 8-0.425\ 73B-0.299\ 64B^2}{1-0.600\ 57B}B^3x_t+\varepsilon_t$$

再考察回归残差序列$\{\varepsilon_t\}$的性质，残差序列的偏自相关图（图 6—4）显示显著的 2 阶截尾性。

```
                     Partial Autocorrelations
Lag    Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  1     0.893 07   |                   . |******************  |
  2    -0.431 03   |           ********* | .                  |
  3    -0.129 87   |                 *** | .                  |
  4     0.023 19   |                   . | .                  |
  5     0.040 54   |                   . |*.                  |
  6    -0.026 57   |                   .*| .                  |
  7    -0.024 96   |                   . | .                  |
  8     0.019 93   |                   . | .                  |
  9     0.000 01   |                   . | .                  |
 10     0.087 21   |                   . |**                  |
```

图 6—4 回归残差序列偏自相关图

所以对残差序列 $AR(2)$再拟合模型，最后得到的拟合模型如下：

$$\begin{cases} y_t = 53.26 + \dfrac{-0.54-0.38B-0.52B^2}{1-0.55B}B^3 x_t + \epsilon_t \\ \epsilon_t = \dfrac{1}{1-1.53B+0.64B^2}a_t \end{cases}$$

式中，$\{a_t\}$ 为零均值白噪声序列。

该输出序列模型的 AIC=8.292 811，SBC=34.006 07。显然这个拟合模型比不考虑输入序列的单纯的 $AR(1,2,4)$疏系数模型优化多了。

该模型拟合效果图如图 6—5 所示。

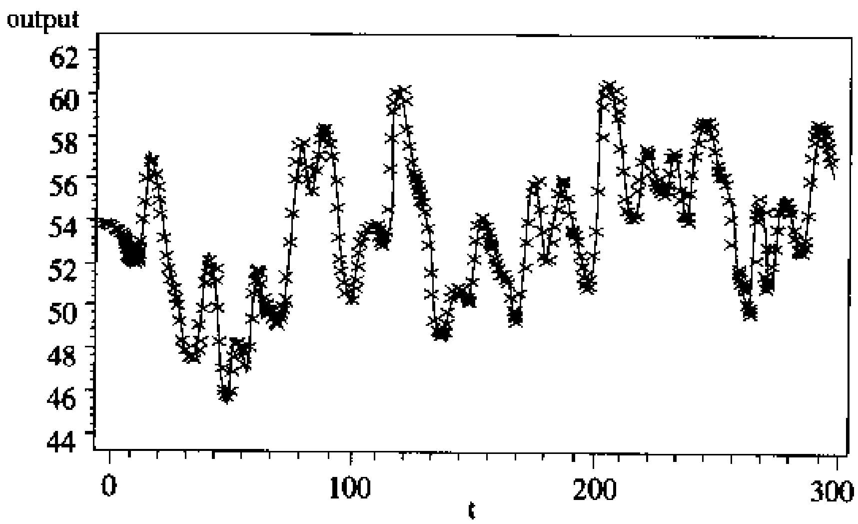


图 6—5 带输入序列模型拟合效果图

图中，星号为序列观察值；曲线为序列拟合值。该图直观显示带输入序列模型拟合效果良好。

6.2 虚假回归

当响应序列 $\{y_t\}$ 和输入变量序列 $\{x_{1t}\}$，$\{x_{2t}\}$，…，$\{x_{kt}\}$ 都平稳时，可以构建带输入变量回归的 $ARIMAX$ 模型来拟合响应序列的变化。

$$y_t = \mu + \sum_{k=1}^{k} \frac{\Theta_i(B)}{\Phi_i(B)}B^{l_i}x_{it} + \frac{\Theta(B)}{\Phi(B)}a_t$$

如果平稳性条件不满足的话，我们就不能大胆地构造 $ARIMAX$ 模型，因为这时容易产生虚值回归的问题。

为了正确理解虚假回归的意义，我们考虑最简单的一元线性动态回归模型：

$$y_t = \beta_0 + \beta_1 x_t + \upsilon_t \tag{6.2}$$

为了检验模型的显著性，我们要对拟合模型进行检验：

$$H_0:\beta_1=0\leftrightarrow H_1:\beta_1\neq 0$$

假定响应序列 $\{y_t\}$ 和输入序列 $\{x_t\}$ 相互独立，那就说明响应序列和输入序列之间没有显著的线性相关关系，理论上，检验结果应该接受 $\beta_1=0$ 的原假设。

假如检验结果恰好和理论结果相反，支持 β_1 显著非零的备择假设，那么我们就会得出响应序列和输入序列之间具有显著线性相关性的错误结论，并接受一个本不应该成立的回归模型（6.2），这就犯了第一类错误（纳伪错误）。

由于样本的随机性，纳伪错误始终都会存在，我们使用显著性水平 α 控制犯纳伪错误的概率：

$$\Pr(H_1\,|\,H_0)=\alpha$$

通常我们采用 t 统计量进行参数显著性检验：

$$t=\frac{\beta_1}{\sigma_\beta}$$

当响应序列和输入序列都平稳时，该统计量服从自由度为样本容量的 t 分布。当 $|t|\leqslant t_{\alpha/2}(n)$时，可以将纳伪错误发生的概率准确地控制在显著性水平 α 以内，即

$$\Pr\{|t|\leqslant t_{\alpha/2}(n)\,|\,\text{平稳序列}\}\leqslant\alpha$$

当响应序列和输入序列不平稳时，随机模拟的结果显示，检验统计量 $t=\frac{\beta_1}{\sigma_\beta}$ 将不再服从 t 分布，这时 t 统计量样本分布的方差远远大于 t 分布的方差，如果仍然采用 t 分布的临界值进行检验，拒绝原假设的概率就会大大增加，即

$$\Pr\{|t|\leqslant t_{\alpha/2}(n)\,|\,\text{非平稳序列}\}\geqslant\alpha$$

这将导致我们无法控制纳伪错误，非常容易接受回归模型显著成立的错误结论，这种现象就称为虚假回归。

6.3 单位根检验

由于虚假回归问题的存在，所以在进行动态回归模型拟合时，必须先检验各序列的平稳性。只有当各序列都平稳时，才可以大胆地使用 $ARIMAX$ 模型拟合多元序列之间的动态回归关系。

在前面我们已经介绍过序列平稳性的图检验方法，由于图检验带有很强的主观色彩，为了客观起见，人们开始研究各种序列平稳性的统计检验方法，其中应

用最广的是单位根检验。

6.3.1 DF 检验

一、DF 统计量

以 1 阶自回归序列为例：

$$x_t=\phi_1 x_{t-1}+\varepsilon_t,\varepsilon_t \overset{i.i.d}{\sim} N(0,\sigma_\varepsilon^2) \tag{6.3}$$

该序列的特征方程为：

$$\lambda-\phi_1=0$$

特征根为：

$$\lambda=\phi_1$$

当特征根在单位圆内时：

$$|\phi_1|<1$$

该序列平稳。

当特征根在单位圆上或单位圆外时：

$$|\phi_1|\geqslant 1$$

该序列非平稳。

所以可以通过检验特征根是在单位圆内还是单位圆上（外），来检验序列的平稳性，这种检验就称为单位根检验。

由于现实生活中绝大多数序列都是非平稳序列，所以单位根检验的原假设定为：

$$H_0:\text{序列 } x_t \text{ 非平稳} \Leftrightarrow H_0:|\phi_1|\geqslant 1$$

相应的备择假设定为：

$$H_0:\text{序列 } x_t \text{ 平稳} \Leftrightarrow H_1:|\phi_1|<1$$

检验统计量为 t 统计量：

$$t(\phi_1)=\frac{\hat{\phi}_1-\phi_1}{S(\hat{\phi}_1)} \tag{6.4}$$

式中，$\hat{\phi}_1$ 为参数 ϕ_1 的最小二乘估计值，$S(\hat{\phi}_1)=\sqrt{\dfrac{S_T^2}{\sum_{t=1}^{T}x_{t-1}^2}}$，$S_T^2=\dfrac{\sum_{t=1}^{T}(x_t-\hat{\phi}_1 x_{t-1})^2}{T-1}$。

当 $\phi_1=0$ 时，$t(\phi_1)$的极限分布为标准正态分布：

$$t(\phi_1)=\frac{\hat{\phi}_1}{S(\hat{\phi}_1)}\xrightarrow{\text{极限}}N(0,1)$$

当 $|\phi_1|<1$ 时，$t(\phi_1)$ 的渐近分布为标准正态分布：

$$t(\phi_1)=\frac{\hat{\phi}_1-\phi_1}{S(\hat{\phi}_1)}\xrightarrow{\text{渐近}}N(0,1)$$

但当 $|\phi_1|=1$ 时，$t(\phi_1)$ 的渐近分布将不再为正态分布。Dickey 和 Fuller 对该检验统计量的样本分布进行了研究。为了区分传统的 t 分布检验统计量，记

$$\tau=\frac{|\hat{\phi}_1|-1}{S(\hat{\phi}_1)}$$

该统计量称为 DF（Dickey-Fuller）检验统计量，它的极限分布为：

$$\frac{\int_0^1 W(r)\mathrm{d}W(r)}{\sqrt{\int_0^1[W(r)]^2\mathrm{d}r}}$$

式中，$W(r)$为自由度为 r 的维纳过程（Weiner process）。所谓维纳过程具有如下性质：

(1)$W(1)\sim N(0,1)$

(2)$\sigma W(r)\sim N(0,\sigma^2 r)$

(3)$[W(r)]^2/r-\chi^2(1)$

随机模拟的结果显示 τ 检验统计量的极限分布为对称钟形分布，和正态分布的形状非常相似，但均值略有偏移。

DF 检验为单边检验，当显著性水平取为 α 时，记 τ_α 为 DF 检验的 α 分位点，则

当 $\tau\leqslant\tau_\alpha$ 时，拒绝原假设，认为序列 x_t 显著平稳；

当 $\tau>\tau_\alpha$ 时，接受原假设，认为序列 x_t 非平稳。

1979 年，Dickey 和 Fuller 使用蒙特卡洛模拟方法算出了 DF 统计量的百分位表，为 DF 检验扫清了最后的技术难题，使 DF 检验成为最常用的单位根检验。

二、DF 检验的等价表达

在式（6.3）等号两边同时减去 x_{t-1}，得到如下等式：

$$x_t-x_{t-1}=(\phi_1-1)x_{t-1}+\varepsilon_t \tag{6.5}$$

记

$$\rho=|\phi_1|-1$$

式（6.5）等价于

$$\nabla x_t=\rho x_{t-1}+\varepsilon_t$$

DF 检验可以通过对参数 ρ 的检验等价进行：

$$H_0:\rho=0 \leftrightarrow H_1:\rho<0$$

相应的 DF 检验统计量为：

$$\tau=\frac{\hat{\rho}}{S(\hat{\rho})}$$

其中，$S(\hat{\rho})$为参数 ρ 的样本标准差。

三、DF 检验的三种类型

在前面的分析中，我们始终以最简单的中心化 1 阶自回归 S 序列为分析重点，实际上，DF 检验还有更为广泛的使用，它可以用于如下三种序列的单位根检验。

第一种类型：无常数均值、无趋势的 1 阶自回归过程。

$$x_t=\phi_1 x_{t-1}+\varepsilon_t,\varepsilon_t \overset{i,i,d}{\sim} N(0,\sigma_\varepsilon^2)$$

这就是我们之前一直分析的类型。

第二种类型：有常数均值、无趋势的 1 阶自回归过程。

$$x_t=\mu+\phi_1 x_{t-1}+\varepsilon_t,\varepsilon_t \overset{i.i.d}{\sim} N(0,\sigma_\varepsilon^2)$$

在这种场合下，通过最小二乘估计方法，可以得到未知参数 μ 和 ϕ_1 的估计值。通过检验特征根的 ϕ_1 性质，可以考察消除常数位移之后的中心化序列 $\{x_t-\mu\}$ 的平稳性：

$$x_t-\mu=\phi_1 x_{t-1}+\varepsilon_t$$

假设条件如下确定：

$H_0:|\phi_1|\geqslant 1$（序列 $\{x_t-\mu\}$ 非平稳）

$H_1:|\phi_1|<1$（序列 $\{x_t-\mu\}$ 平稳）

第三种类型：既有常数均值，又有线性趋势的 1 阶自回归过程。

$$x_t=\mu+\beta t+\phi_1 x_{t-1}+\varepsilon_t,\varepsilon_t \overset{i,i,d}{\sim} N(0,\sigma_\varepsilon^2)$$

在这种场合下，通过最小二乘估计方法，可以得到未知参数 μ，β 和 ϕ_1 的估计值。通过检验特征根 ϕ_1 的性质，可以考察消除了常数位移和线性趋势之后的中心化 1 阶自回归序列 $\{x_t-\mu-\beta t\}$ 的平稳性。

$$x_t-\mu-\beta t=\phi_1 x_{t-1}+\varepsilon_t$$

假设条件如下确定：

$H_0:|\phi_1|\geqslant 1$（序列 $\{x_t-\mu-\beta t\}$ 非平稳）

$H_1:|\phi_1|<1$（序列 $\{x_t-\mu-\beta t\}$ 平稳）

【例 6.2】 对 1978—2002 年中国农村居民家庭人均纯收入对数序列 $\{\ln x_t\}$

和生活消费支出对数序列 $\{\ln y_t\}$ 进行 DF 检验（数据见附录 1.21）。

1. 直观判断

对中国农村居民家庭人均收入对数序列 $\{\ln x_t\}$ 和家庭人均生活消费支出对数序列 $\{\ln y_t\}$ 绘制时序图，如图 6—6 所示。

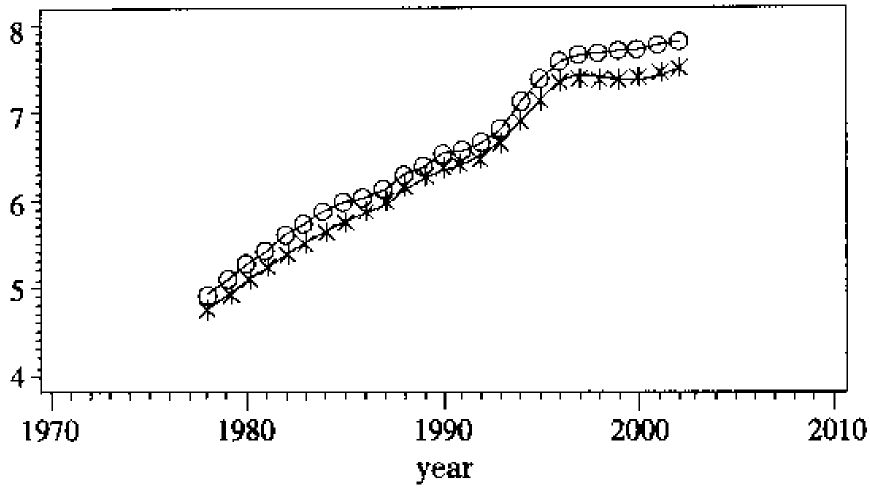


图 6—6 中国农村居民家庭人均收入与生活消费支出对数序列时序图

图中，圆形曲线代表中国农村居民家庭人均收入对数序列 $\{\ln x_t\}$；星号曲线代表中国农村居民家庭人均生活消费支出对数序列 $\{\ln y_t\}$。时序图显示这两个对数序列显著非平稳。

2. 对中国农村居民家庭人均纯收入对数序列 $\{\ln x_t\}$ 进行 DF 检验

单位根检验结果如表 6—1 所示。($\alpha=0.05$)

表 6—1

类型	延迟阶数	模型结构	τ 检验统计量的值	$\Pr<\tau$
类型 1：无均值、无趋势模型	0	$\ln x_t=\varepsilon_t$	7.05	0.999 9
	1	$\ln x_t=\frac{\varepsilon_t}{1-\phi_1 B}$	1.09	0.922 4
类型 2：有均值、无趋势模型	0	$\ln x_t=\mu+\varepsilon_t$	−2.18	0.219 0
	1	$\ln x_t=\mu+\frac{\varepsilon_t}{1-\phi_1 B}$	−1.16	0.673 0
类型 3：有趋势模型	0	$\ln x_t=\mu+\beta t+\varepsilon_t$	−0.7	0.961 7
	1	$\ln x_t=\mu+\beta t+\frac{\varepsilon_t}{1-\phi_1 B}$	−3.14	0.121 5

检验结果显示，无论考虑何种类型的模型，τ 检验统计量的 P 值均显著大于 α ($\alpha=0.05$)，所以可以认为中国农村居民家庭人均纯收入对数序列 $\{\ln x_t\}$ 显

著非平稳。

3. 对中国农村居民家庭人均生活消费支出对数序列 $\{\ln y_t\}$ 进行 DF 检验如表 6—2 所示。(α=0.05)

表 6—2

类型	延迟阶数	τ 检验统计量的值	$Pr<\tau$
类型 1	0	6.54	0.999 9
	1	1.08	0.921 7
类型 2	0	−2.27	0.188 7
	1	−1.41	0.559 8
类型 3	0	−0.46	0.978 4
	1	−3.31	0.089 2

同收入序列的检验结果一样，在显著性水平 α 取为 0.05 时，可以认为中国农村居民家庭人均生活消费支出对数序列 $\{\ln(y_t)\}$ 非平稳。

显然，这两个序列的 DF 检验结果与根据时序图得到的直观判断完全一致。

6.3.2 ADF 检验

DF 检验只适用于 1 阶自回归过程的平稳性检验，但是实际上绝大多数时间序列不会是一个简单的 $AR(1)$ 过程。为了使 DF 检验能适用于 $AR(p)$ 过程的平稳性检验，人们对 DF 检验进行了一定的修正，得到增广 DF 检验（augmented Dickey-Fuller），简记为 ADF 检验。

一、ADF 检验的原理

对任一 $AR(p)$ 过程

$$x_t=\phi_1 x_{t-1}+\cdots+\phi_p x_{t-p}+\epsilon_t \tag{6.6}$$

它的特征方程为：

$$\lambda^p-\phi_1\lambda^{p-1}-\cdots-\phi_p=0$$

如果该方程所有的特征根都在单位圆内，即

$$|\lambda_i|<1,\ i=1,\ 2,\ \cdots,\ p$$

则序列 $\{x_t\}$ 平稳。

如果有一个特征根存在，不妨设

$$\lambda_1=1$$

则序列 $\{x_t\}$ 非平稳，且自回归系数之和恰好等于 1：

$$\lambda^p-\phi_1\lambda^{p-1}-\cdots-\phi_p=0\overset{\lambda=1}{\Rightarrow}1-\phi_1-\cdots-\phi_p=0\Rightarrow\phi_1+\phi_2+\cdots+\phi_p=1$$

因而，对于 $AR(p)$ 过程我们可以通过检验自回归系数之和是否等于 1 来考察该序列的平稳性。

为了便于检验，我们对式（6.6）进行等价变换：

$$\begin{aligned} x_t - x_{t-1} &= \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} - x_{t-1} + \varepsilon_t \\ &= (\phi_2 + \cdots + \phi_p) x_{t-1} + \phi_1 x_{t-1} - x_{t-1} - (\phi_2 + \cdots + \phi_p) x_{t-1} \\ &\quad + \phi_2 x_{t-2} + (\phi_3 + \cdots + \phi_p) x_{t-2} - (\phi_2 + \cdots + \phi_p) x_{t-2} \\ &\quad + \phi_3 x_{t-3} + (\phi_4 + \cdots + \phi_p) x_{t-3} + \cdots - \phi_p x_{t-p+1} + \phi_p x_{t-p} + \varepsilon_t \end{aligned}$$

整理上式，得

$$\nabla x_t = (\phi_1 + \cdots + \phi_p - 1) x_{t-1} - (\phi_3 + \cdots + \phi_p) \nabla x_{t-1} - \cdots - \phi_p \nabla x_{t-p+1}$$

简记为：

$$\nabla x_t = \rho x_{t-1} + \beta_1 \nabla x_{t-1} + \cdots + \beta_p \nabla x_{t-p} + \varepsilon_t \qquad (6.7)$$

式中：

$$\rho = \phi_1 + \phi_2 + \cdots + \phi_p - 1$$

$$\beta_j = -\phi_{j+1} - \phi_{j+2} - \cdots - \phi_p,\ j = 1,\ 2,\ \cdots,\ p-1$$

若序列 $\{x_t\}$ 平稳，则

$$\phi_1 + \phi_2 + \cdots + \phi_p < 1$$

等价于

$$\rho < 0$$

若序列 $\{x_t\}$ 非平稳，则至少存在一个单位根，有

$$\phi_1 + \phi_2 + \cdots + \phi_p = 1$$

等价于

$$\rho = 0$$

则 $AR(p)$ 过程单位根检验的假设条件可以如下确定：

$$H_0: \rho = 0 \text{(序列 } x_t \text{ 非平稳)} \leftrightarrow H_1: \rho < 0 \text{ (序列 } x_t \text{ 平稳)}$$

构造 ADF 检验统计量：

$$\tau = \frac{\hat{\rho}}{S(\hat{\rho})}$$

式中，$S(\hat{\rho})$ 为参数 ρ 的样本标准差。

通过蒙特卡洛方法，可以得到 τ 检验统计量的临界假表。显然 DF 检验是 ADF 检验在自相关阶数为 1 时的一个特例，所以它们统称为 ADF 检验。

二、ADF 检验的三种类型

和 DF 检验一样，ADF 检验也可以用于如下三种类型的单位根检验。

第一种类型：无常数均值、无趋势的 p 阶自回归过程。

$$x_t=\phi_1 x_{t-1}+\cdots+\phi_p x_{t-p}+\varepsilon_t$$

第二种类型：有常数均值、无趋势的 p 阶自回归过程。

$$x_t=\mu+\phi_1 x_{t-1}+\cdots+\phi_p x_{t-p}+\varepsilon_t$$

第三种类型：既有常数均值，又有线性趋势的 p 阶自回归过程。

$$x_t=\mu+\beta t+\phi_1 x_{t-1}+\cdots+\phi_p x_{t-p}+\varepsilon_t$$

【例 6.2 续】 对 1978—2002 年中国农村居民家庭人均纯收入对数差分后序列 $\{\nabla\ln x_t\}$ 和生活消费支出对数差分后序列 $\{\nabla\ln y_t\}$ 进行 ADF 检验。

1. 中国农村居民家庭人均纯收入对数差分后序列 $\{\nabla\ln x_t\}$ ADF 检验结果（3 阶延迟）。如表 6—3 所示。

表 6—3

类型	延迟阶数	τ 检验统计量的值	$\Pr<\tau$
类型 1	0	−1.38	0.149 8
	1	−1.37	0.152 3
	2	−1.36	0.155 5
	3	−1.34	0.160 6
类型 2	0	−1.92	0.319 6
	1	−2.23	0.203 3
	2	−3.1	0.042 1
	3	−1.9	0.327 8
类型 3	0	−2.08	0.531 1
	1	−2.41	0.364 6
	2	−3.37	0.082 0
	3	−1.99	0.573 2

上述检验结果显示，只有如下模型能显著拒绝原假设，如表 6—4 所示。($\alpha=0.05$)

表 6—4

模型	$\Pr<\tau$
$\nabla\ln x_t=\mu+\frac{\varepsilon_t}{1-\phi_1 B-\phi_2 B^2}$	0.042 1

这说明，对数差分后的纯收入序列是有常数均值的平稳序列，该平稳序列 2 阶自相关。

2. 中国农村居民家庭人均生活消费支出对数差分后序列 $\{\nabla\ln y_t\}$ ADF 检验结果（3 阶延迟），如表 6—5 所示。

表 6—5

类型	延迟阶数	τ 检验统计量的值	Pr<τ
类型 1	0	−1.27	0.181 3
	1	−1.8	0.068 8
	2	−1.3	0.169 9
	3	−1.06	0.251 4
类型 2	0	−1.89	0.329 7
	1	−3.35	0.024 6
	2	−2.81	0.073 7
	3	−1.72	0.408 6
类型 3	0	−2.16	0.489 7
	1	−3.62	0.051 1
	2	−3.32	0.090 2
	3	−2.21	0.461 3

检验结果显示，只有如下模型能显著拒绝原假设，如表 6—6 所示。($\alpha=0.05$)

表 6—6

模型	Pr<τ
$\nabla \ln y_t=\mu+\frac{\varepsilon_t}{1-\phi_1 B}$	0.024 6

这说明，对数差分后的生活消费支出序列是具有常数均值的平稳序列，该平稳序列 1 阶自相关。

6.3.3 PP 检验

使用 ADF 检验有一个基本假定：

$$\mathrm{Var}(\varepsilon_t)=\sigma_\varepsilon^2$$

这导致 ADF 检验主要适用于方差齐性场合，它对于异方差序列的平稳性检验效果不佳。

Phillips 和 Perron 于 1988 年对 ADF 检验进行了非参数修正，提出了 Phillips-Perron 检验统计量。该检验统计量既可适用于异方差场合的平稳性检验，又服从相应的 ADF 检验统计量的极限分布。

使用 Phillips-Perron 检验（简记为 PP 检验），残差序列 $\{\varepsilon_t\}$ 需要满足如下三个条件。

1. 均值恒为零

$$E(\varepsilon_t)=0$$

2. 方差及至少一个高阶矩存在

$$\sup_t E(|\varepsilon_t|^2)<\infty$$

且对于某个 $\beta>2$，$\sup_t E(|\varepsilon_t|^\beta)<\infty$。

由于没有假定 $E(|\varepsilon_t|^2)$常数，所以这个条件实际上意味着允许异方差性存在。

3. 非退化极限分布存在

$$\sigma_S^2=\lim_{T\to\infty}E(T^{-1}S_T^2)\text{存在且为正值}$$

式中，T 为序列长度；$S_T=\sum_{t=1}^{T}\varepsilon_t$。

以 1 阶自回归模型

$$x_t=\phi_1 x_{t-1}+\varepsilon_t$$

为例，介绍 PP 检验的构造原理。

假设 $\hat{\phi}_1$ 是 ϕ_1 的最小二乘估计值（OLS），那么 $\hat{\phi}_1$ 的方差通常可以定义为：

$$\sigma^2=\lim_{T\to\infty}T^{-1}\sum_{t=1}^{T}E(\varepsilon_t^2)$$

当 $\{\varepsilon_t\}$ 为白噪声序列时，有

$$\sigma^2=\sigma_S^2 \tag{6.8}$$

式中，$\sigma_S^2=\lim_{T\to\infty}E(T^{-1}S_T^2)$，$S_T=\sum_{t=1}^{T}\varepsilon_t$。

如果 $\{\varepsilon_t\}$ 不满足白噪声条件，那么方差等式（6.8）将不再成立。

为了直观说明式（6.8）不成立，不妨假设 $\{\varepsilon_t\}$ 服从 MA（1）过程，即

$$\varepsilon_t=\upsilon_t-\theta_1\upsilon_{t-1},\upsilon_t\overset{i,i,d}{\sim}N(0,\sigma_\upsilon^2)$$

那么

$$\sigma^2=E(\varepsilon_t^2)=(1+\theta_1^2)\sigma_\upsilon^2$$

而

$$\sigma_S^2=E(\varepsilon_1^2)+2\sum_{j=2}^{\infty}E(\varepsilon_1\varepsilon_j)=(1+\theta_1^2-2\theta_1)\sigma_\upsilon^2=(1-\theta_1)^2\sigma^2\upsilon$$

显然　$\sigma^2\neq\sigma_S^2$

Phillips 和 Perron 正是利用这种不等性，利用 σ^2 和 σ_S^2 的估计值对 ADF 检验的 τ 统计量进行了非参数修正，修正后的统计量如下：

$$Z(\tau)=\tau(\hat{\sigma}^2/\hat{\sigma}_{Sl}^2)-(1/2)(\hat{\sigma}_{Sl}^2-\hat{\sigma}^2)T\sqrt{\hat{\sigma}_{Sl}^2\sum_{t=2}^{T}(x_{t-1}-\bar{x}_{T-1})^2}$$

式中：

(1)$\hat{\sigma}^2$ 是 σ^2 的样本估计值，即

$$\hat{\sigma}^2 = T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_t^2$$

(2)假设可以估计$\{\varepsilon_t\}$显著自相关的延迟阶数为 l，$\hat{\sigma}_{Sl}^2$是σ_S^2 的样本估计值：

$$\hat{\sigma}_{Sl}^2 = T^{-}\sum_{t=1}^{l}\hat{\varepsilon}_t^2 + 2T^{-1}\sum_{j=1}^{t}w_j(l)\sum_{t=j+1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_{t-j}$$

式中，$w_j(l)=1-\frac{1}{l+1}$，这个权重确保了 $\hat{\sigma}_{Sl}^2$ 为正值。

(3) $\bar{x}_{T-1} = \frac{1}{T-1}\sum_{t=1}^{T-1}x_t$

在单位根检验原假设成立的条件下：

$H_0:\phi_1=1$ (序列$\{x_t\}$非平稳)

修正后的 $Z(\tau)$统计量和 τ 统计量具有相同的极限分布。这就意味着对于异方差序列只需要在原来 τ 统计量的基础上进行一定的修正，构造出 $Z(\tau)$统计量。$Z(\tau)$统计量不仅考虑到自相关误差所产生的影响，还可以继续使用 τ 统计量的临界值表进行检验，而不需要拟合新的临界值表。

【例 6.2 续】 对 1978—2002 年中国农村居民家庭人均纯收入对数差分后序列 $\{\nabla\ln x_t\}$ 和生活消费支出对数差分后序列 $\{\nabla\ln y_t\}$ 进行 PP 检验。

1. 中国农村居民家庭人均纯收入对数差分后序列 $\{\nabla\ln x_t\}$ PP 检验结果（3 阶延迟）如表 6—7 所示。

表 6—7

类型	延迟阶数	τ 检验统计量的值	$Pr<\tau$
类型 1	0	−1.38	0.149 8
	1	−1.4	0.144 7
	2	−1.43	0.138 1
	3	−1.39	0.148 4
类型 2	0	−1.92	0.319 6
	1	−2.05	0.266 1
	2	−2.16	0.224 0
	3	−2.06	0.262 0
类型 3	0	−2.08	0.531 1
	1	−2.22	0.455 9
	2	−2.34	0.399 5
	3	−2.22	0.454 8

检验结果显示，考虑到异方差性，上述三种类型均不平稳（$\alpha=0.05$）。

2. 中国农村居民家庭人均生活消费支出对数差分后序列 $\{\nabla \ln y_t\}$ PP 检验结果（3 阶延迟）如表 6—8 所示。

表 6—8

类型	延迟阶数	τ 检验统计量的值	$Pr<\tau$
类型 1	0	−1.27	0.181 3
	1	−1.37	0.152 4
	2	−1.38	0.151 3
	3	−1.3	0.171 4
类型 2	0	−1.89	0.329 7
	1	−2.17	0.222 8
	2	−2.2	0.211 0
	3	−2.06	0.260 3
类型 3	0	−2.16	0.489 7
	1	−2.43	0.353 6
	2	−2.47	0.339 7
	3	−2.31	0.413 7

和收入序列的检验结果一样，考虑到异方差性，上述三种类型均不平稳。（$\alpha=0.05$）

3. 对中国农村居民家庭人均纯收入对数差分后序列 $\{\nabla \ln x_t\}$ 和人均生活消费支出对数差分后序列 $\{\nabla \ln y_t\}$ 再进行一次差分，对 2 阶差分后序列 $\{\nabla^2 \ln x_t\}$ 与 $\{\nabla^2 \ln y_t\}$ 进行 PP 检验（3 阶延迟），检验结果如表 6—9 所示。

表 6—9

类型	延迟阶数	$\{\nabla^2 \ln x_t\}$		$\{\nabla^2 \ln y_t\}$	
		τ 检验统计量的值	$Pr<\tau$	τ 检验统计量的值	$Pr<\tau$
类型 1	0	−4.29	0.000 2	−3.42	0.001 6
	1	−4.29	0.000 2	−3.47	0.001 4
	2	−4.31	0.000 2	−3.44	0.001 5
	3	−4.27	0.000 2	−3.34	0.001 9
类型 2	0	−4.23	0.003 5	−3.37	0.023 7
	1	−4.23	0.003 5	−3.42	0.021 0
	2	−4.25	0.003 3	−3.39	0.022 4
	3	−4.21	0.003 7	−3.28	0.028 3
类型 3	0	−4.12	0.019 3	−3.27	0.098 3
	1	−4.12	0.019 4	−3.33	0.088 1
	2	−4.15	0.018 4	−3.29	0.093 4
	3	−4.09	0.020 5	−3.17	0.115 8

检验结果显示，2 阶差分后的对数收入序列 $\{\nabla^2\ln x_t\}$ 与 2 阶差分后的对数消费序列 $\{\nabla^2\ln y_t\}$ 均显著平稳。

6.4 协　整

6.4.1 单整与协整

一、单整（integration）的概念

在单位根检验的过程中，如果检验结果显著拒绝原假设，即说明序列 $\{x_t\}$ 显著平稳，不存在单位根，这时称序列 $\{x_t\}$ 为零阶**单整**序列，简记为 $x_t \sim I(0)$。

假如原假设不能被显著拒绝，说明序列 $\{x_t\}$ 为非平稳序列，存在单位根。这时可以考虑对该序列进行适当阶数的差分，以消除单位根实现平稳。

假如原序列 1 阶差分后平稳，说明原序列存在一个单位根，这时称原序列为 1 阶单整序列，简记为 $x_t \sim I(1)$。

假如原序列至少需要进行 d 阶差分才能实现平稳，说明原序列存在 d 个单位根，这时称原序列为 d 阶单整序列，简记为 $x_t \sim I(d)$。

考虑例 6.2 的检验结果，如果不考虑异方差性的影响，使用 ADF 检验，则 1 阶差分后的对数序列 $\{\nabla\ln x_t\}$，$\{\nabla\ln y_t\}$ 显著平稳，即 $\ln x_t \sim I(1)$，$\ln y_t \sim I(1)$。

如果考虑异方差性的影响，使用 PP 检验，则 2 阶差分后的对数序列 $\{\nabla^2\ln x_t\}$，$\{\nabla^2\ln y_t\}$ 显著平稳，即 $\ln x_t \sim I(2)$，$\ln y_t \sim I(2)$。

二、单整序列的性质

单整衡量的是单个序列的平稳性，它具有如下重要性质：

1. 若 $x_t \sim I(0)$，对于任意非零实数 a，b，有

$$a + bx_t \sim I(0)$$

2. 若 $x_t \sim I(d)$，对于任意非零实数 a，b，有

$$a + bx_t \sim I(d)$$

3. 若 $x_t \sim I(0)$，$y_t \sim I(0)$，对于任意非零实数 a，b，有

$$z_t = ax_t + by_t \sim I(0)$$

4. 若 $x_t \sim I(d)$，$y_t \sim I(c)$，对于任意非零实数 a，b，有

$$z_t = ax_t + by_t \sim I(k)$$

式中，$k \leqslant \max[d, c]$。

三、协整（cointegration）的概念

在现实生活中我们会发现，有些序列自身的变化虽然是非平稳的，但是序列与序列之间却具有非常密切的长期均衡关系。

比如说例 6.2 所分析的农村家庭人均纯收入对数序列 $\{\ln x_t\}$ 和人均生活消费支出对数序列 $\{\ln y_t\}$，单整分析显示这两个序列都是非平稳的。但是将这两个序列联合起来考虑，通过观察它们的时序图（图 6—6），我们却发现它们之间具有非常稳定的线性相关关系。当收入增多时，生活消费支出也增多，它们的变化速度几乎一致。这种稳定的同变关系让我们怀疑它们之间具有一种内在的平稳机制，导致了它们自身的变化虽然是不平稳的，但是彼此之间却具有长期均衡关系。

为了有效地衡量序列之间是否具有长期均衡关系，Engle 和 Granger 于 1987 年提出了协整的概念。

假定自变量序列为 $\{x_1\}$，…，$\{x_k\}$，响应变量序列为 $\{y_t\}$，构造回归模型

$$y_t=\beta_0+\sum_{k=1}^{k}\beta_i x_{it}+\varepsilon_t$$

假定回归残差序列 $\{\varepsilon_t\}$ 平稳，我们称响应序列 $\{y_t\}$ 与自变量序列 $\{x_1\}$，…，$\{x_k\}$ 之间具有协整关系。

协整概念的提出有着非常重要的意义，因为我们之前一直不敢大胆地对非平稳序列构建动态回归模型，是担心非平稳序列容易产生虚假回归的问题。而虚假回归之所以会产生是因为残差序列不平稳。如果非平稳序列之间具有协整关系，那就说明残差序列平稳，那就不会产生虚假回归问题了。

这说明要将输入变量引入响应序列建模，不一定要所有的序列都平稳，而只需要它们之间具有协整关系。这个限制条件显然比 Cox 和 Jenkins 要求所有序列都平稳的限制条件要宽松许多，这极大地拓宽了动态回归模型的适用范围。

6.4.2 协整检验

多元非平稳序列之间能否建立动态回归模型，关键在于它们之间是否具有协整关系。所以要对多元非平稳序列建模必须得先进行协整检验，也称为 Engle-Granger 检验，简称为 EG 检验。

一、假设条件

由于自然界中绝大多数序列之间不具有协整关系，所以 EG 检验的假设条件如下确定：

H_0：多元非平稳序列之间不存在协整关系

H_1：多元非平稳序列之间存在协整关系

由于协整关系主要是通过考察回归残差的平稳性确定，所以上述假设条件等价于：

H_0：回归残差序列 $\{\varepsilon_t\}$ 非平稳

H_1：回归残差序列 $\{\varepsilon_t\}$ 平稳

二、检验步骤

EG 检验也称为 EG 两步法，它遵循如下两个步骤进行。

步骤一：建立响应序列与输入序列之间的回归模型：

$$y_t=\hat{\beta}_0+\hat{\beta}_1x_{1t}+\cdots+\hat{\beta}_kx_{kt}+\varepsilon_t$$

式中，$\hat{\beta}_0$，$\hat{\beta}_1$，…，$\hat{\beta}_k$ 是最小二乘估计值。

步骤二：对回归残差序列 $\{\varepsilon_t\}$ 进行平稳性检验。

我们主要是采用单位根检验的方法来考察回归残差序列的平稳性，所以，假设条件等价于

$$H_0:\varepsilon_t\sim I(k),k\geqslant1\leftrightarrow H_1:\varepsilon_t\sim I(0)$$

EG 检验的原理与计算公式和 DF 检验的原理与计算公式相同，但是蒙特卡洛模拟的结果显示它们的临界值略有不同。EG 检验的临界值不仅与位移项、趋势项等因素有关，而且还与回归模型中非平稳变量的个数相关。Mackinnon 提供了 EG 检验的临界值表，并将 EG 检验的临界值表与 ADF 检验的临界值表结合在一起。当非平稳序列的个数为 1 时（$N=1$），对应的就是 ADF 检验；当非平稳序列的个数大于等于 2 时（$N\geqslant2$），对应的就是 EG 检验。

【例 6.2 续】 对 1978—2002 年中国农村居民家庭人均纯收入对数序列 $\{\ln x_t\}$ 和生活消费支出对数序列 $\{\ln y_t\}$ 进行 EG 检验。

第一步：构造回归模型。

利用最小二乘估计方法，构造出回归模型如下：

$$\ln y_t=0.968\ 32\ln x_t+\varepsilon_t$$

第二步：残差序列单位根检验。

表 6—8 检验结果如表 6—10 所示。

表 6—10

类型	延迟阶数	τ 检验统计量的值	$\Pr<\tau$
类型 1	0	−1.33	0.162 9
	1	−1.69	0.084 5
	2	−1.93	0.052 8
类型 2	0	−1.28	0.622 2
	1	−1.64	0.445 5
	2	−1.85	0.348 4

根据类型 1 延迟 1 阶的检验结果，我们可以以 91.55%（1－0.084 5）的把握断定残差序列平稳且具有 1 阶自相关性：

$$\varepsilon_t=\phi_1\varepsilon_{t-1}+\upsilon_t,\upsilon_t\overset{i,i,d}{\sim}N(0,\sigma_\upsilon^2)$$

也就是说我们可以以 91.55%的把握认为中国农村居民家庭人均纯收入对数序列 $\ln\{x_t\}$ 和生活消费支出对数序列 $\ln\{y_t\}$ 之间存在协整关系，并可以构建如下动态回归模型：

$$\ln y_t=\beta\ln x_t+\frac{\upsilon_t}{1-\phi_1 B},\upsilon_t\overset{i,i,d}{\sim}N(0,\sigma_\upsilon^2)$$

利用条件最小二乘估计方法，得到该模型的口径为：

$$\ln y_t=0.968\ 21\ln x_t+\frac{\upsilon_t}{1-0.838\ 13B},\upsilon_t\overset{i,i,d}{\sim}N(0,0.000\ 893)$$

检验结果显示回归模型显著成立，参数显著非零，残差序列 $\{\upsilon_t\}$ 为白噪声序列。

上述分析说明，尽管中国农村居民家庭人均纯收入对数序列 $\ln(x_t)$ 和生活消费支出对数序列 $\ln\{y_t\}$ 都是非平稳序列，但是由于它们之间具有协整关系，所以可以建立动态回归模型准确地拟合它们之间的互动关系。

对对数序列 $\ln\{y_t\}$ 和该序列的拟合值序列同时进行指数运算，可以得到中国农村家庭生活消费支出拟合效果图如图 6—7 所示。

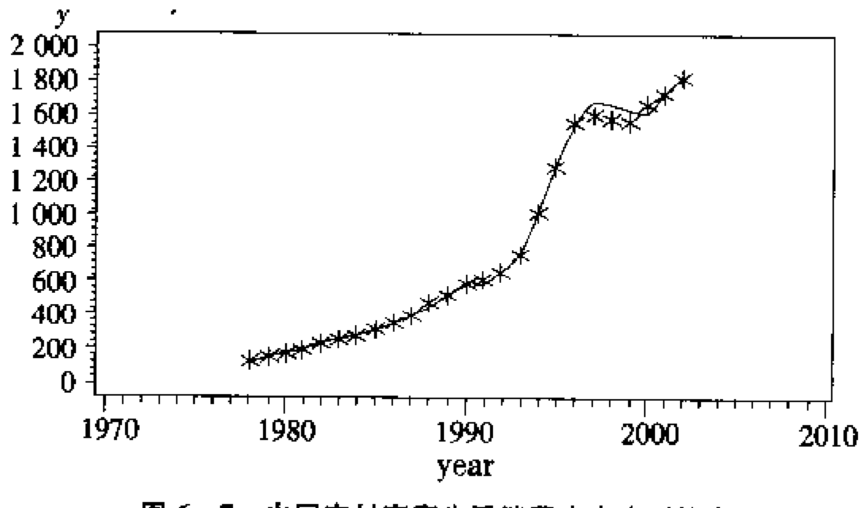


图 6—7　中国农村家庭生活消费支出序列拟合图

图中，星号为中国农村家庭生活消费支出观察值；曲线为模型拟合值。

6.5 误差修正模型

误差修正模型（error correction model）简称为 ECM，最初由 Hendry 和 Anderson 于 1977 年提出，它常常作为协整回归模型的补充模型出现。由协整模型度量序列之间的长期均衡关系，而 ECM 模型则解释序列的短期波动关系。

误差修正模型的构造原理如下：

假设非平稳响应序列 $\{y_t\}$ 与非平稳输入序列 $\{x_t\}$ 之间具有协整关系，即

$$y_t = \beta x_t + \varepsilon_t \tag{6.9}$$

则回归残差序列为平稳序列：

$$\varepsilon_t = y_t - \beta_t \sim I(0)$$

在式（6.9）等号两边同时减去 y_{t-1}，则有

$$y_t - y_{t-1} = \beta x_t - y_{t-1} + \varepsilon_t \tag{6.10}$$

将 $y_{t-1} = \beta x_{t-1} + \varepsilon_{t-1}$ 带入式（6.10）等号右边，得

$$y_t - y_{t-1} = \beta x_t - \beta x_{t-1} - \varepsilon_{t-1} + \varepsilon_t \tag{6.11}$$

假定 β 的最小二乘估计值为 $\hat{\beta}$，则 $\hat{\varepsilon}_{t-1} = y_{t-1} - \hat{\beta} x_{t-1}$ 代表的是上一期的误差，特别记作 ECM_{t-1}，则式（6.11）可以整理成如下形式：

$$\nabla y_t = \beta \nabla x_t - ECM_{t-1} + \varepsilon_t \tag{6.12}$$

这说明响应序列的当期波动（∇y_t）主要会受到三方面的短期波动的影响：

1. 输入序列的当期波动 ∇x_t；
2. 上一期的误差 ECM_{t-1}；
3. 纯随机波动 ε_t。

为了定量地测定这三方面影响的大小，尤其是为了测定上期误差 ECM_{t-1} 对当期波动 ∇y_t 的影响，我们可以构建 ECM 模型，模型结构如下：

$$\nabla y_t = \beta_0 \nabla x_t + \beta_1 ECM_{t-1} + \varepsilon_t$$

式中，β_1 称为误差修正系数，表示误差修正项对当期波动的修正力度。根据误差修正模型的推导原理（式（6.12）），可以确定 $\beta_1 < 0$，即误差修正机制是一个负反馈机制。

以例 6.2 的应用背景对误差修正模型的负反馈机制进行直观解释，当 $ECM_{t-1} < 0$ 时，等价于 $y_{t-1} > \hat{\beta} x_{t-1}$，即上期真实支出比估计支出大，这种误差反馈回来，会导致下期支出适当压缩，即 $\nabla y_t < 0$。

反之，$ECM_{t-1}>0$，等价于 $y_{t-1}<\hat{\beta}x_{t-1}$，即上期真实支出比估计支出小，这种误差反馈回来，会导致下期支出适当增加，即$\nabla y_t>0$。

【例 6.2 续】 对 1978—2002 年中国农村居民家庭人均纯收人对数序列$\{\ln x_t\}$和生活消费支出对数序列$\{\ln y_t\}$构造 ECM 模型。

在上一节中，我们已经通过 EG 检验证明中国农村居民家庭人均纯收人对数序列 $\ln\{x_t\}$和生活消费支出对数序列 $\ln\{y_t\}$具有协整关系，即

$$\ln y_t=0.968\ 32\ln x_t+\varepsilon_t$$

这个协整回归模型揭示了中国农村居民家庭人均生活消费支出与人均纯收人之间的长期均衡关系。

为了研究生活消费支出的短期波动特征，我们利用差分序列 $\{\nabla\ln y_t\}$，$\{\nabla\ln x_t\}$和前期误差序列 $\{ECM_{t-1}\}$ 构建 ECM 模型：

$$\nabla\ln y_t=\beta_0\ \nabla\ln x_t+\beta_1 ECM_{t-1}+\varepsilon_1$$

本例中 $ECM_{t-1}=\ln y_{t-1}-0.968\ 32\ln x_{t-1}$。

计算结果显示 ECM 模型为：

$$\nabla\ln y_t=0.957\ 9\ \nabla\ln x_t-0.153\ 7ECM_{t-1}+\varepsilon_t$$

方程检验结果和参数检验结果如表 6—11 所示。

表 6—11

方程检验		参数检验		
R^2	0.954 3	参数	t 值	$Pr>\|t\|$
DW	1.493 4	β_0	20.81	<0.000 1
$Pr<DW$	0.073 8	β_1	1.18	0.249 9

方程检验结果显示该方程显著线性相关。参数检验结果显示收人的当期波动对生活消费支出的当期波动有显著性影响（β_0 显著），但上期误差（ECM）对当期波动的影响不显著（β_1 不显著）。而且从回归系数的绝对值大小可以看出收人的当期波动对生活消费支出的当期波动调整幅度很大，每增加 1 元的收入会增加 0.957 9 元的生活消费支出，但上期误差（ECM）对生活消费支出的当期波动调整幅度不大，单位调整比例为$-0.153\ 7$。

6.6 习 题

1. 单位根检验的原理是什么？使用单位根检验方法重新考察例 2.1 至例 2.5 中序列的平稳性，并将检验结果与图检验方法得出的结果进行比较研究。

2. 某地区过去 38 年谷物产量序列如表 6--12 所示。

表 6—12

24.5	33.7	27.9	27.5	21.7	31.9	36.8	29.9	30.2	32.0	34.0
19.4	36.0	30.2	32.4	36.4	36.9	31.5	30.5	32.3	34.9	30.1
36.9	26.8	30.5	33.3	29.7	35.0	29.9	35.2	38.3	35.2	35.5
36.7	26.8	38.0	31.7	32.6						

这些年该地区相应的降雨量序列如表 6—13 所示。

表 6—13

9.6	12.9	9.9	8.7	6.8	12.5	13.0	10.1	10.1	10.1	10.8
7.8	16.2	14.1	10.6	10.0	11.5	13.6	12.1	12.0	9.3	7.7
11.0	6.9	9.5	16.5	9.3	9.4	8.7	9.5	11.6	12.1	8.0
10.7	13.9	11.3	11.6	10.4						

(1) 使用单位根检验，分别考察这两个模型的平稳性。

(2) 选择适当模型，分别拟合这两个序列的发展。

(3) 确定这两个序列之间是否具有协整关系。

(4) 如果这两个序列之间具有协整关系，请建立适当的模型拟合谷物产量序列的发展。

3. 在一定浓度的溶液中（CC=0.5），考察草履虫（aramecium aurelia）和某种草履虫掠食动物（didinium nasutum）之间的动态数量变化，相关数据如表 6—14 所示。

表 6—14

时间 day	被掠食者 ind/ml	掠食者 ind/ml	时间 day	被掠食者 ind/ml	掠食者 ind/ml	时间 day	被掠食者 ind/ml	掠食者 ind/ml
0.00	15.65	5.76	12.00	27.46	65.40	24.00	121.70	17.82
0.50	53.57	9.05	12.50	41.46	51.35	24.50	185.20	26.04
1.00	73.34	17.26	13.00	44.73	28.24	25.00	175.30	65.61
1.50	93.93	41.97	13.50	88.42	23.27	25.50	139.00	76.30
2.00	115.40	55.97	14.00	105.70	38.09	26.00	77.11	96.07
2.50	76.57	74.91	14.50	155.20	14.97	26.50	57.29	68.84
3.00	32.83	62.52	15.00	205.50	24.84	27.00	54.79	54.79
3.50	23.74	27.04	15.50	312.70	49.56	27.50	75.38	35.80
4.00	56.70	18.77	16.00	213.70	75.93	28.00	87.73	32.48
4.50	86.37	31.11	16.50	163.40	104.00	28.50	136.40	24.21

续前表

时间 day	被掠食者 ind/ml	掠食者 ind/ml	时间 day	被掠食者 ind/ml	掠食者 ind/ml	时间 day	被掠食者 ind/ml	掠食者 ind/ml
5.00	121.00	58.31	17.00	85.78	106.40	29.00	290.60	35.73
5.50	71.48	73.13	17.50	48.64	100.60	29.50	345.80	55.50
6.00	55.78	63.21	18.00	44.49	84.08	30.00	271.60	93.41
6.50	31.84	52.46	18.50	63.44	45.30	30.50	156.10	117.30
7.00	26.87	40.07	19.00	71.66	35.37	31.00	71.10	95.02
7.50	53.24	27.67	19.50	127.70	35.35	31.50	43.86	85.92
8.00	65.59	26.00	20.00	206.90	41.10	32.00	30.64	82.60
8.50	81.23	24.32	20.50	309.90	52.62	32.50	35.56	66.08
9.00	143.90	21.00	21.00	156.50	120.20	33.00	52.03	63.58
9.50	237.90	33.35	21.50	63.30	112.80	33.50	37.99	37.99
10.00	276.60	64.67	22.00	77.29	92.14	34.00	62.71	25.60
10.50	222.20	94.34	22.50	45.11	65.72	34.50	103.90	23.10
11.00	137.20	103.40	23.00	57.45	33.54	35.00	187.20	37.09
11.50	46.45	82.74	23.50	69.80	21.14			

(1) 考察这两个生物之间的动态关系，检验它们是否具有协整关系。

(2) 选择适当的模型拟合这两个生物之间的动态互动关系，并预测未来一周这两个生物的浓度。

4. 收集中国城市居民的平均收入序列和中国农村居民的平均收入序列，考察这两个序列各自具有怎样的特征，它们之间是否具有协整关系?

5. 收集改革开放以来中国的进出口序列，考察中国的进口和出口之间具有怎样的相关关系，有无协整关系存在?

6. 从《中国统计年鉴》中寻找两个具有协整关系的序列，并解释它们之间具有稳定相关的现实意义。

6.7 上机指导

SAS 系统中的 *ARIMA* 过程可以支持单位根检验并能建立带输入变量的 *ARIMAX* 模型。以如下临时数据集 example6 _ 1 为例，介绍单位根检验与 *ARIMAX* 模型建模命令。

一、建立数据集，并查看时序图

```
data example6 _ 1;
input x y@@;
```

```
t= _n_;
cards;
-2.94   9.83  -2.14  12.63   1.01  14.77
 2.84  17.29  -0.79  18.07   1.46  17.38
 5.44  19.17   1.65   9.12   6.53  22.82
 8.93  23.58   8.67  15.19   8.36  22.43
 9.79  17.83  11.67  25.49   9.70  28.40
 9.18  23.15  11.13  19.70   9.39  22.32
12.89  30.01   8.45  21.27   6.66  11.52
 4.15  15.57   2.57   9.91   2.29  23.28
-3.28  13.75  -5.21   3.38  -3.74  15.81
-8.73  12.41 -15.89   5.54 -12.15   4.83
-10.86 14.79 -17.16   4.14 -18.55  -5.36
-11.42  4.79 -16.02   0.91 -14.36  -5.49
-17.98  6.01 -16.94   2.78 -17.52  -2.49
-13.44 10.30 -14.11  -0.32 -15.16   2.35
;
proc gplot;
plot x*t=1 y*t=2/overlay;
symbol1 c=black i=join v=none;
symbol2 c=red i=join v=none w=2 l=2;
run;
```

输出时序图如图 6—8 所示。

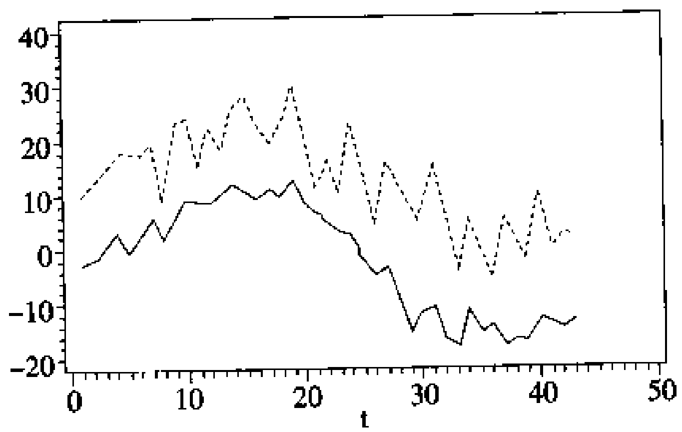


图 6—8　x 序列和 y 序列的联合时序图

图中，实线为 x 序列时序图；虚线为 y 序列时序图。从时序图中可以看出 x 序列、y 序列均显著非平稳，这个直观判断我们还可以通过单位根检验验证。同时时序图显示这两个序列具有某种同变关系，可以考虑建立 ARIMAX 模型。

二、单位根检验

使用如下命令可以对 x 序列和 y 序列进行单位根检验：

```
proc arima data=example6_1;
identify var=x stationarity= (adf=1);
identify var=y stationarity= (adf=1);
run;
```

语句说明：

"identify var=x stationarity= (adf=1);"，在 ARIMA 过程 IDENTIFY 命令中有一个平稳性检验的选项 stationarity=，该选项支持 ADF 检验、PP 检验和 RW 检验，本例指定进行 1 阶自相关 ADF 检验。

序列 x 的单位根检验输出结果如图 6—9 所示。

Augmented Dickey-Fuller Unit Root Tests							
Type	Lags	Rho	Pr <Rho	Tau	Pr <Tau	F	Pr> F
Zero Mean	0	−0.968 0	0.473 4	−0.48	0.500 3		
	1	−0.372 1	0.593 6	−0.21	0.603 1		
Sing Mean	0	−1.233 8	0.857 8	−0.60	0.860 2	0.35	0.981 5
	1	−0.600 2	0.915 0	−0.34	0.909 6	0.32	0.986 1
Trend	0	−6.564 2	0.671 2	−2.28	0.434 6	3.34	0.529 0
	1	−5.628 8	0.752 7	−2.24	0.457 2	3.50	0.498 1

图 6—9 序列 x 的单位根检验结果

第一列输出的是检验的模型类型，第二列输出的是自相关延迟阶数，这两列联合确定了检验模型的具体形式，如表 6—15 所示。

表 6—15

Type	Lags	模型具体结构
Zero Mean	0	$x_t=\varepsilon_t,\ \varepsilon_t \overset{i.i.d}{\sim} N(0,\sigma_\varepsilon^2)$
	1	$x_t=\phi_1 x_{t-1}+\varepsilon_t,\ \varepsilon_t \overset{i.i.d}{\sim} N(0,\sigma_\varepsilon^2)$
Single Mean	0	$x_t=\mu+\varepsilon_t,\ \varepsilon_t \overset{i.i.d}{\sim} N(0,\sigma_\varepsilon^2)$
	1	$x_t=\mu+\phi_1 x_{t-1}+\varepsilon_t,\ \varepsilon_t \overset{i.i.d}{\sim} N(0,\sigma_\varepsilon^2)$
Trend	0	$x_t=\mu+\beta t+\varepsilon_t,\ \varepsilon_t \overset{i.i.d}{\sim} N(0,\sigma_\varepsilon^2)$
	1	$x_t=\mu+\beta t+\phi_1 x_{t-1}+\varepsilon_t,\ \varepsilon_t \overset{i.i.d}{\sim} N(0,\sigma_\varepsilon^2)$

第三列、第四列输出的是 Rho 统计量的值及检验 P 值。

第五列、第六列输出的是 Tau 统计量的值及检验 P 值。关于 Tau 统计量的构造原理及检验标准我们在第 6 章中详细介绍过。本例中 Tau 统计量的 P 值显著大于 0.05，可以认为序列显著非平稳，至少存在一个单位根。

第七列、第八列输出的是回归模型显著性检验 F 统计量的值及检验 P 值。

序列 y 的单位根检验输出结果如图 6—10 所示。

Augmented Dickey-fuller Unit Root Tests

Type	Lags	Rho	Pr ＜Rho	Tau	Pr ＜Tau	F	Pr＞F
Zero Mean	0	−4.840 8	0.123 9	−1.62	0.098 9		
	1	−2.345 3	0.288 4	−1.13	0.230 2		
Sing Mean	0	−13.813 2	0.037 2	−2.74	0.076 4	3.76	0.144 6
	1	−6.881 7	6.258 6	−1.64	0.451 9	1.40	0.718 5
Trend	0	−26.207 7	0.005 7	−4.47	0.004 9	10.19	0.001 0
	1	−20.954 3	0.027 5	−3.41	0.064 6	6.12	0.077 4

图 6—10　序列 y 的单位根检验结果

根据第五列、第六列输出的结果我们可以判断当显著性水平 α 取为 0.05 时，序列 y 非平稳，但消除线性趋势之后序列平稳。

三、*ARIMAX* 模型建模

两个非平稳序列建立回归模型容易产生虚值回归的问题，除非它们之间具有协整关系。ARIMA 过程支持直观的协整判断，并能进行 *ARIMAX* 模型阶数识别，相关命令如下：

```
proc arima;
identify var=y crosscorr=x;
estimate method=ml input=x noint plot;
```

语句说明：

(1) “identify var=y crosscorr=x;”，该语句是指定序列 y 为因变量，序列 x 为输入序列，识别序列 y 的性质，并考察它与序列 x 之间的相关性。

该命令输出结果除了 IDENTIFY 常规输出的描述统计信息、样本自相关图、逆自相关图、偏自相关图及白噪声检验结果之外，还会附加一项序列 y 各阶延迟序列与序列 x 之间的相关图，如图 6—11 所示。

```
                              Crosscorrelations
Lag   Covariance    Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
-10   11.358 791     0.121 77   |                 .  |**  .                |
 -9   23.173 398     0.248 42   |                 .  |*****.               |
 -8   29.670 969     0.318 08   |                 .  |******               |
 -7   40.891 559     0.438 37   |                 .  |*********            |
 -6   49.008 605     0.525 38   |                 .  |***********          |
 -5   54.433 302     0.583 54   |                 .  |************         |
 -4   62.623 700     0.671 34   |                 .  |*************        |
 -3   71.193 567     0.763 21   |                 .  |***************      |
 -2   68.546 999     0.734 84   |                 .  |***************      |
 -1   71.869 487     0.770 45   |                 .  |***************      |
  0   81.884 714     0.877 82   |                 .  |******************   |
  1   75.853 355     0.813 16   |                 .  |****************     |
  2   69.122 812     0.741 01   |                 .  |***************      |
  3   70.504 831     0.755 82   |                 .  |***************      |
  4   64.907 600     0.695 82   |                 .  |**************       |
  5   57.572 151     0.617 18   |                 .  |************         |
  6   51.451 107     0.551 57   |                 .  |***********          |
  7   43.492 929     0.466 25   |                 .  |*********            |
  8   34.846 268     0.373 56   |                 .  |*******              |
  9   27.282 295     0.292 47   |                 .  |******               |
 10   19.649 211     0.210 64   |                 .  |****  .              |

                      "."marks two standard errors
```

图 6—11　序列 y 与序列 x 之间的相关图

图中第一列为因变量序列的延迟阶数，第二列为延迟序列与输入序列之间的协方差，第三列为相关系数，后面是相关图。

本例相关图显示序列 y 在延迟阶数为零时与序列 x 相关关系最大。因此可以将序列 y 与序列 x 同期建模。

假如显示结果是序列 y 在延迟阶数为 $-k$ 时与序列 x 相关关系最大，说明序列 y 会受到序列 x 滞后 k 期的影响，因此在建模时应该将序列 y 与序列 x 的 k 期延迟序列 lag$k(x)$建立回归模型。

(2) "estimate method=ml input=x plot;"，该语句是指令系统建立以序列 y 为因变量，序列 x 为自变量的回归模型，并输出残差序列的相关图，以便于我们检验残差序列的平稳性及相关性。

假如是要建立序列 y 与序列 x 的 k 期延迟序列 lag$k(x)$之间的回归模型，相关命令为：

estimate method=ml input=（k $ x) plot;

执行 ESTIMATE 命令会输出 8 部分内容：

(1) 参数估计值；

(2) 拟合统计量；

(3) 参数的协方差阵；

(4) 残差白噪声检验结果；

(5) 残差自相关图；

(6) 残差逆自相关图；

(7) 残差偏自相关图；

(8) 模型拟合结果。

本例输出的残差序列相关图如图 6—12 所示。

```
                         Autocorrelations Plot of Residuals

Lag Covariance Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1 Std Error
  0  19.805 880  1.000 00 |                    |********************|         0
  1  -2.554 549  -.128 98 |              .  ***|     .              | 0.154 303
  2  -4.569 046  -.230 69 |              .*****|     .              | 0.156 849
  3   3.273 938  0.165 30 |             .      |***   .             | 0.164 730
  4  -2.689 905  -.135 81 |             .   ***|      .             | 0.168 633
  5  -0.082 542  -.004 17 |             .      |      .             | 0.171 217
  6  -1.609 092  -.081 24 |             .    **|      .             | 0.171 220
  7  -3.475 644  -.175 49 |             .  ****|      .             | 0.172 135
  8  -3.555 474  -.179 52 |             .  ****|      .             | 0.176 343
  9   2.579 873  0.130 26 |             .      |***   .             | 0.180 642
 10   3.713 389  -.187 49 |             .      |****  .             | 0.182 865

                         "."marks two standard errors
```

图 6—12　残差序列自相关图

自相关图显示所有延迟自相关系数都在 2 倍标准差范围内，可以认为残差序列平稳。

如果对自己的直观判断没有把握的话，我们还可以输出残差序列，并对残差序列进行单位根检验，以验证自己的直观判断。在原程序中添加如下命令可以实现这个功能：

```
forecast lead=0 id=t out=out;
proc arima data=out;
identify var=residual stationarity= (adf=2);
run;
```

语句说明：

(1) “forecast lead=0 id=t out=out;”，该语句是为了输出估计结果，并将拟合结果存在临时数据集 OUT 中，其中该数据集最后一列数据就是拟合残差 RESIDUAL。

(2) “identify var=residual stationarity= (adf=2);”，对残差序列进行 2 阶自相关单位根检验，检验结果显示残差序列显著平稳。如图 6—13 所示。

Augmented Dickey-Fuller Unit Root Tests							
Type	Lags	Rho	Pr <Rho	Tau	Pr <Tau	F	Pr> F
Zero Mean	0	−46.289 9	<.000 1	−7.23	<.000 1		
	1	−76.674 9	<.000 1	−6.00	<.000 1		
	2	−49.933 4	<.000 1	−3.75	0.000 4		
Sing Mean	0	−46.291 6	0.000 3	−7.14	0.000 1	25.48	0.0000
	1	−76.815 7	0.000 3	−5.93	0.000 2	17.58	0.0010
	2	−50.103 1	0.000 3	−3.70	0.007 6	6.86	0.0048
Trend	0	−46.284 3	<.000 1	−7.04	<.000 1	24.83	0.0010
	1	−76.786 3	<.000 1	−5.84	<.000 1	17.10	0.0010
	2	−50.052 3	<.000 1	−3.65	0.038 5	6.67	0.0528

图 6—13 残差序列单位根检验结果

残差序列平稳，说明序列 y 与序列 x 之间具有协整关系，我们可以大胆地在这两个序列之间建立回归模型而不必担心虚假回归问题。

再返回 ESTIMATE 命令相关输出结果，考察残差序列白噪声检验结果。如图 6—14 所示。

Autocorrelation Check of Residuals									
To Lag	ChiSquare	DF	Pr> ChiSq	Autocorrelations					
6	5.74	6	0.452 9	−0.129	−0.231	0.165	−0.136	−0.004	−0.081
12	15.87	12	0.197 4	−0.175	−0.180	0.130	0.187	−0.241	0.072
18	28.57	18	0.053 9	0.289	−0.151	0.190	−0.049	−0.207	0.032
24	32.05	24	0.125 8	0.010	−0.058	−0.079	0.154	−0.023	−0.063

图 6—14 残差序列白噪声检验结果

输出结果显示延迟各阶 LB 统计量的 P 值都大于显著性水平 0.05，可以认为残差序列为白噪声序列，结束分析。

根据最后输出的模型拟合结果可知最后的拟合模型口径为：

$$y_t=14.591\,57+0.773\,671+\varepsilon_t,\quad \varepsilon_t\overset{i.i.d}{\sim}N\ (0,\ 19.805\,88)$$

假如残差序列白噪声检验结果为非白噪声序列，就可以考察残差自相关图、偏自相关图、逆自相关图的性质，为残差序列构造合适的 $ARMA(p, q)$ 模型，在原程序中添加相关命令如下：

```
estimate p=p q=q input=x;
```

输出拟合模型结构如下：

$$y_t=\mu+\beta x+\frac{1-\theta_1 B-\cdots-\theta_q B^q}{1-\phi_1 B-\cdots-\phi_q B^q}\varepsilon_t$$

一旦拟合模型显著，我们还可以利用拟合模型预测序列未来走势并绘制拟合效果图，相关命令形如：

```
forecast lead=5 id=t out=result;
proc gplot data=result;
plot y * t=1 forecast * t=2 l95 * t=3 u95 * t=3/overlay;
symbol1 c=black i=none v=star;
symbol2 c=rd i=join v=none;
symbol3 c=green i=join v=none;
run;
```

21世纪

附录 1

附录 1.1　　1770—1869 年太阳黑子年度数据

年份	黑子数	年份	黑子数	年份	黑子数	年份	黑子数
1820	16	1833	8	1846	62	1859	94
1821	7	1834	13	1847	98	1860	96
1822	4	1835	57	1848	124	1861	77
1823	2	1836	122	1849	96	1862	59
1824	8	1837	138	1850	66	1863	44
1825	17	1838	103	1851	64	1864	47
1826	36	1839	86	1852	54	1865	30
1827	50	1840	63	1853	39	1866	16
1828	62	1841	37	1854	21	1867	7
1829	67	1842	24	1855	7	1868	37
1830	71	1843	11	1856	4	1869	74
1831	48	1844	15	1857	23		
1832	28	1845	40	1858	55		

资料来源：George E. P. Box，Gwilym M. Jenkins，Gregory C. Reinsel，*Time Series Analysis Forecasting and Control*（Third Edition），Prentice-Hall，Inc.

附录 1.2 1964—1999 年中国纱年产量序列 单位：万吨

年份	纱产量	年份	纱产量
1964	97.0	1982	335.4
1965	130.0	1983	327.0
1966	156.5	1984	321.9
1967	135.2	1985	353.5
1968	137.7	1986	397.8
1969	180.5	1987	436.8
1970	205.2	1988	465.7
1971	190.0	1989	476.7
1972	188.6	1990	462.6
1973	196.7	1991	460.8
1974	180.3	1992	501.8
1975	210.8	1993	501.5
1976	196.0	1994	489.5
1977	223.0	1995	542.3
1978	238.2	1996	512.2
1979	263.5	1997	559.8
1980	292.6	1998	542.0
1981	317.0	1999	567.0

资料来源：北京市统计局计算中心：《北京五十年》，北京，北京电子出版物出版中心，1999。

附录 1.3 1962 年 1 月—1975 年 12 月平均每头奶牛月产奶量序列 单位：磅

589	561	640	656	727	697	640	599	568	577
553	582	600	566	653	673	742	716	660	617
583	587	565	598	628	618	688	705	770	736
678	639	604	611	594	634	658	622	709	722
782	756	702	653	615	621	602	635	677	635
736	755	811	798	735	697	661	667	645	688
713	667	762	784	837	817	767	722	681	687
660	698	717	696	775	796	858	826	783	740
701	706	677	711	734	690	785	805	871	845
801	764	725	723	690	734	750	707	807	824
886	859	819	783	740	747	711	751	804	756
860	878	942	913	869	834	790	800	763	800
826	799	890	900	961	935	894	855	809	810
766	805	821	773	883	898	957	924	881	837
784	791	760	802	828	778	889	902	969	947
908	867	815	812	773	813	834	782	892	903
966	937	896	858	817	827	797	843		

资料来源：Cryer (1986).

附录 1.4　　1949—1998 年北京市每年最高气温序列　　单位：摄氏度

年份	温度	年份	温度
1949	38.8	1974	35.8
1950	35.6	1975	38.4
1951	38.3	1976	35.0
1952	39.6	1977	34.1
1953	37.0	1978	37.5
1954	33.4	1979	35.9
1955	39.6	1980	35.1
1956	34.6	1981	38.1
1957	36.2	1982	37.3
1958	37.6	1983	37.2
1959	36.8	1984	36.1
1960	38.1	1985	35.1
1961	40.6	1986	38.5
1962	37.1	1987	36.1
1963	39.0	1988	38.1
1964	37.5	1989	35.8
1965	38.5	1990	37.5
1966	37.5	1991	35.7
1967	35.8	1992	37.5
1968	40.1	1993	35.8
1969	35.9	1994	37.2
1970	35.3	1995	35.0
1971	35.2	1996	36.0
1972	39.5	1997	38.2
1973	37.5	1998	37.2

资料来源：北京市统计局计算中心：《北京五十年》，北京，北京电子出版物出版中心，1999。

附录 1.5　　1950—1998 年北京市城乡居民定期储蓄所占比例序列（%）

年份	定期储蓄	年份	定期储蓄
1950	83.5	1975	83.8
1951	63.1	1976	84.5
1952	71.0	1977	84.8
1953	76.3	1978	83.9
1954	70.5	1979	83.9
1955	80.5	1980	81.0
1956	73.6	1981	82.2

续前表

年份	定期储蓄	年份	定期储蓄
1957	75.2	1982	82.7
1958	69.1	1983	82.3
1959	71.4	1984	80.9
1960	73.6	1985	80.3
1961	78.8	1986	81.3
1962	84.4	1987	81.6
1963	84.1	1988	83.4
1964	83.3	1989	88.2
1965	83.1	1990	89.6
1966	81.6	1991	90.1
1967	81.4	1992	88.2
1968	84.0	1993	87.0
1969	82.9	1994	87.0
1970	83.5	1995	88.3
1971	83.2	1996	87.8
1972	82.2	1997	84.7
1973	83.2	1998	80.2
1974	83.5		

资料来源：北京市统计局计算中心：《北京五十年》，北京，北京电子出版物出版中心，1999。

附录 1.6　　某加油站连续 57 天的 OVERSHORTS

78	−58	53	−63	13	−6	−16	−14
3	−74	89	−48	−14	32	56	−86
−66	50	26	59	−47	−83	2	−1
124	−106	113	−76	−47	−32	39	−30
6	−73	18	2	−24	23	−38	91
−56	−58	1	14	−4	77	−127	97
10	−28	−17	23	−2	48	−131	65
−17							

资料来源：Brockwell and Davis.

附录 1.7　　1971 年 9 月—1993 年 6 月澳大利亚季度常住人口变动序列　　单位：千人

63.2	67.9	55.8	49.5	50.2	55.4
49.9	45.3	48.1	61.7	55.2	53.1
49.5	59.9	30.6	30.4	33.8	42.1
35.8	28.4	32.9	44.1	45.5	36.6
39.5	49.8	48.8	29.0	37.3	34.2
47.6	37.3	39.2	47.6	43.9	49.0
51.2	60.8	67.0	48.9	65.4	65.4
67.6	62.5	55.1	49.6	57.3	47.3
45.5	44.5	48.0	47.9	49.1	48.8
59.4	51.6	51.4	60.9	60.9	56.8
58.6	62.1	64.0	60.3	64.6	71.0
79.4	59.9	83.4	75.4	80.2	55.9
58.5	65.2	69.5	59.1	21.5	62.5
170.0	−47.4	62.2	60.0	33.1	35.3
43.4	42.7	58.4	34.4		

资料来源：根据 ABS. Compare CIVPOP. DAT. June 1971 -June 1993 提供的 1971 年 6 月—1993 年 6 月澳大利亚季度常住人口估计值序列计算所得。

附录 1.8　　连续读取 70 个化学反应数据

47	64	23	71	38	64	55	41	59	48	71	35	57	40
58	44	80	55	37	74	51	57	50	60	45	57	50	45
25	59	50	71	56	74	50	58	45	54	36	54	48	55
45	57	50	62	44	64	43	52	38	59	55	41	53	49
34	35	54	45	68	38	50	60	39	59	40	57	54	23

资料来源：O'Donovan，Consec. Readings Batch Chemical Process，R. B. Miller et al.

附录 1.9　　上海证券交易所 1991 年 1 月—2001 年 10 月每月末上证指数 x_t　　单位：点

130.44	133.47	120.19	113.94	114.83	137.56	143.80	178.43	180.92
218.60	259.60	292.75	313.24	364.66	381.24	445.38	1 234.71	1 191.19
1 052.07	823.27	702.32	507.25	724.60	780.39	1 198.48	1 339.88	925.91
1 358.78	935.48	1 007.05	881.07	895.68	890.27	814.82	984.93	833.80
770.25	770.98	704.46	592.56	556.26	469.29	333.92	785.33	791.15
654.98	683.59	647.87	562.59	549.26	646.92	579.93	700.51	630.58
695.55	723.87	722.43	717.32	641.13	555.29	537.34	552.93	556.39
681.16	643.65	804.25	822.48	809.94	875.52	976.71	1 032.95	917.02
964.74	1 040.27	1 234.62	1 393.75	1 285.18	1 250.27	1 189.76	1 221.06	1 097.38
1 180.39	1 139.63	1 194.10	1 222.91	1 206.53	1 243.01	1 343.44	1 411.20	1 339.20

续前表

1 316.91	1 150.22	1 242.09	1 217.31	1 247.42	1 146.70	1 134.67	1 090.09	1 158.05
1 120.92	1 279.32	1 689.42	1 601.45	1 627.12	1 570.70	1 504.56	1 434.97	1 366.58
1 535.00	1 714.58	1 800.22	1 836.32	1 894.55	1 928.11	2 023.54	2 021.20	1 910.16
1 961.29	2 070.61	2 073.48	2 065.61	1 959.18	2 112.78	2 119.18	2 213.18	2 218.03
1 920.32	1 834.14	1 764.87	1 689.17					

附录 1.10　　北京市 1995—2000 年月平均气温序列　　单位：摄氏度

月＼年	1995	1996	1997	1998	1999	2000
1	−0.7	−2.2	−3.8	−3.9	−1.6	−6.4
2	2.1	−0.4	1.3	2.4	2.2	−1.5
3	7.7	6.2	8.7	7.6	4.8	8.1
4	14.7	14.3	14.5	15.0	14.4	14.6
5	19.8	21.6	20.0	19.9	19.5	20.4
6	24.3	25.4	24.6	23.6	25.4	26.7
7	25.9	25.5	28.2	26.5	28.1	29.6
8	25.4	23.9	26.6	25.1	25.6	25.7
9	19.0	20.7	18.6	22.2	20.9	21.8
10	14.5	12.8	14.0	14.8	13.0	12.6
11	7.7	4.2	5.4	4.0	5.9	3.0
12	−0.4	0.9	−1.5	0.1	−0.6	−0.6

附录 1.11　　1993—2000 年中国社会消费品零售总额序列　　单位：亿元

	1993	1994	1995	1996	1997	1998	1999	2000
1月	977.5	1 192.2	1 602.2	1 909.1	2 288.5	2 549.5	2 662.1	2 774.7
2月	892.5	1 162.7	1 491.5	1 911.2	2 213.5	2 306.4	2 538.4	2 805.0
3月	942.3	1 167.5	1 533.3	1 860.1	2 130.9	2 279.7	2 403.1	2 627.0
4月	941.3	1 170.4	1 548.7	1 854.8	2 100.5	2 252.7	2 356.8	2 572.0
5月	962.2	1 213.7	1 585.4	1 898.3	2 108.2	2 265.2	2 364.0	2 637.0
6月	1 005.7	1 281.1	1 639.7	1 966.0	2 164.7	2 326.0	2 428.8	2 645.0
7月	963.8	1 251.5	1 623.6	1 888.7	2 102.5	2 286.1	2 380.3	2 597.0
8月	959.8	1 286.0	1 637.1	1 916.4	2 104.4	2 314.6	2 410.9	2 636.0
9月	1 023.3	1 396.2	1 756.0	2 083.5	2 239.6	2 443.1	2 604.3	2 854.0
10月	1 051.1	1 444.1	1 818.0	2 148.3	2 348.0	2 536.0	2 743.9	3 029.0
11月	1 102.0	1 553.8	1 935.2	2 290.1	2 454.9	2 652.2	2 781.5	3 108.0
12月	1 415.5	1 932.2	2 389.5	2 848.6	2 881.7	3 131.4	3 405.7	3 680.0

资料来源：中国经济信息网。

附录 1.12　　1950—1999 年北京市民用车辆拥有量序列　　单位：万辆

年份	车辆拥有量	年份	车辆拥有量
1950	5.43	1975	91.71
1951	6.19	1976	106.70
1952	6.63	1977	119.93
1953	7.18	1978	135.84
1954	8.95	1979	155.49
1955	10.14	1980	178.29
1956	11.74	1981	199.14
1957	12.60	1982	215.75
1958	17.26	1983	232.63
1959	21.07	1984	260.41
1960	22.38	1985	321.12
1961	24.00	1986	361.95
1962	24.80	1987	408.07
1963	26.13	1988	464.38
1964	27.61	1989	511.32
1965	29.95	1990	551.36
1966	33.92	1991	606.11
1967	33.21	1992	691.74
1968	34.80	1993	817.58
1969	37.16	1994	941.95
1970	42.41	1995	1 040.00
1971	49.44	1996	1 100.08
1972	57.74	1997	1 219.09
1973	67.27	1998	1 319.30
1974	78.57	1999	1 452.94

资料来源：北京市统计局计算中心：《北京五十年》，北京，北京电子出版物出版中心，1999。

附录 1.13　　1962 年 1 月—1975 年 12 月平均每头奶牛月产奶量序列　　单位：磅

589	561	640	656	727	697	640	599
568	577	553	582	600	566	653	673
742	716	660	617	583	587	565	598
628	618	688	705	770	736	678	639
604	611	594	634	658	622	709	722
782	756	702	653	615	621	602	635
677	635	736	755	811	798	735	697
661	667	645	688	713	667	762	784

续前表

837	817	767	722	681	687	660	698
717	696	775	796	858	826	783	740
701	706	677	711	734	690	785	805
871	845	801	764	725	723	690	734
750	707	807	824	886	859	819	783
740	747	711	751	804	756	860	878
942	913	869	834	790	800	763	800
826	799	890	900	961	935	894	855
809	810	766	805	821	773	883	898
957	924	881	837	784	791	760	802
828	778	889	902	969	947	908	867
815	812	773	813	834	782	892	903
966	937	896	858	817	827	797	843

资料来源：Cryer (1986).

附录 1.14　　1952—1988 年中国农业实际国民收入指标序列
（以 1952 年农业国民收入总额为基数 100）

年份	农业	年份	农业
1952	100.0	1971	142.0
1953	101.6	1972	140.5
1954	103.3	1973	153.1
1955	111.5	1974	159.2
1956	116.5	1975	162.3
1957	120.1	1976	159.1
1958	120.3	1977	155.1
1959	100.6	1978	161.2
1960	83.6	1979	171.5
1961	84.7	1980	168.4
1962	88.7	1981	180.4
1963	98.9	1982	201.6
1964	111.9	1983	218.7
1965	122.9	1984	247.0
1966	131.9	1985	253.7
1967	134.2	1986	261.4
1968	131.6	1987	273.2
1969	132.2	1988	279.4
1970	139.8		

资料来源：http：//www.cs.few.eur.nl/few/people/franses.

附录 1.15　　　　1917—1975 年美国 23 岁妇女每万人生育率序列

年	每万人生育率	年	每万人生育率
1917	183.1	1947	212.0
1918	183.9	1948	200.4
1919	163.1	1949	201.8
1920	179.5	1950	200.7
1921	181.4	1951	215.6
1922	173.4	1952	222.5
1923	167.6	1953	231.5
1924	177.4	1954	237.9
1925	171.7	1955	244.0
1926	170.1	1956	259.4
1927	163.7	1957	268.8
1928	151.9	1958	264.3
1929	145.4	1959	264.5
1930	145.0	1960	268.1
1931	138.9	1961	264.0
1932	131.5	1962	252.8
1933	125.7	1963	240.0
1934	129.5	1964	229.1
1935	129.6	1965	204.8
1936	129.5	1966	193.3
1937	132.2	1967	179.0
1938	134.1	1968	178.1
1939	132.1	1969	181.1
1940	137.4	1970	165.6
1941	148.1	1971	159.8
1942	174.1	1972	136.1
1943	174.7	1973	126.3
1944	156.7	1974	123.3
1945	143.3	1975	118.5
1946	189.7		

数据来源：Hipel and Mcleod (1994).

附录 1.16　　　　1962—1991 年德国工人季度失业率序列　　　　行数据（%）

1.1	0.5	0.4	0.7	1.6	0.6	0.5	0.7
1.3	0.6	0.5	0.7	1.2	0.5	0.4	0.6

续前表

0.9	0.5	0.5	1.1	2.9	2.1	1.7	2.0
2.7	1.3	0.9	1.0	1.6	0.6	0.5	0.7
1.1	0.5	0.5	0.6	1.2	0.7	0.7	1.0
1.5	1.0	0.9	1.1	1.5	1.0	1.0	1.6
2.6	2.1	2.3	3.6	5.0	4.5	4.5	4.9
5.7	4.3	4.0	4.4	5.2	4.3	4.2	4.5
5.2	4.1	3.9	4.1	4.8	3.5	3.4	3.5
4.2	3.4	3.6	4.3	5.5	4.8	5.4	6.5
8.0	7.0	7.4	8.5	10.1	8.9	8.8	9.0
10.0	8.7	8.8	8.9	10.4	8.9	8.9	9.0
10.2	8.6	8.4	8.4	9.9	8.5	8.6	8.7
9.8	8.6	8.4	8.2	8.8	7.6	7.5	7.6
8.1	7.1	6.9	6.6	6.8	6.0	6.2	6.2

资料来源：*Time Series Models for Business and Economic Forecasting*，Cambridge University，1998.

附录 1.17　　1948—1981 年美国女性（大于 20 岁）月度失业率序列

446	650	592	561	491	592	604	635	580	510
553	554	628	708	629	724	820	865	1 007	1 025
955	889	965	878	1 103	1 092	978	823	827	928
838	720	756	658	838	684	779	754	794	681
658	644	622	588	720	670	746	616	646	678
552	560	578	514	541	576	522	530	564	442
520	484	538	454	404	424	432	458	556	506
633	708	1 013	1 031	1 101	1 061	1 048	1 005	987	1 006
1 075	854	1 008	777	982	894	795	799	781	776
761	839	842	811	843	753	848	756	848	828
857	838	986	847	801	739	865	767	941	846
768	709	798	831	833	798	806	771	951	799
1 156	1 332	1 276	1 373	1 325	1 326	1 314	1 343	1 225	1 133
1 075	1 023	1 266	1 237	1 180	1 046	1 010	1 010	1 046	985
971	1 037	1 026	947	1 097	1 018	1 054	978	955	1 067
1 132	1 092	1 019	1 110	1 262	1 174	1 391	1 533	1 479	1 411
1 370	1 486	1 451	1 309	1 316	1 319	1 233	1 113	1 363	1 245
1 205	1 084	1 048	1 131	1 138	1 271	1 244	1 139	1 205	1 030
1 300	1 319	1 198	1 147	1 140	1 216	1 200	1 271	1 254	1 203
1 272	1 073	1 375	1 400	1 322	1 214	1 096	1 198	1 132	1 193
1 163	1 120	1 164	966	1 154	1 306	1 123	1 033	940	1 151

续前表

1 013	1 105	1 011	963	1 040	838	1 012	963	888	840
880	939	868	1 001	956	966	896	843	1 180	1 103
1 044	972	897	1 103	1 056	1 055	1 287	1 231	1 076	929
1 105	1 127	988	903	845	1 020	994	1 036	1 050	977
956	818	1 031	1 061	964	967	867	1 058	987	1 119
1 202	1 097	994	840	1 086	1 238	1 264	1 171	1 206	1 303
1 393	1 463	1 601	1 495	1 561	1 404	1 705	1 739	1 667	1 599
1 516	1 625	1 629	1 809	1 831	1 665	1 659	1 457	1 707	1 607
1 616	1 522	1 585	1 657	1 717	1 789	1 814	1 698	1 481	1 330
1 646	1 596	1 496	1 386	1 302	1 524	1 547	1 632	1 668	1 421
1 475	1 396	1 706	1 715	1 586	1 477	1 500	1 648	1 745	1 856
2 067	1 856	2 104	2 061	2 809	2 783	2 748	2 642	2 628	2 714
2 699	2 776	2 795	2 673	2 558	2 394	2 784	2 751	2 521	2 372
2 202	2 469	2 686	2 815	2 831	2 661	2 590	2 383	2 670	2 771
2 628	2 381	2 224	2 556	2 512	2 690	2 726	2 493	2 544	2 232
2 494	2 315	2 217	2 100	2 116	2 319	2 491	2 432	2 470	2 191
2 241	2 117	2 370	2 392	2 255	2 077	2 047	2 255	2 233	2 539
2 394	2 341	2 231	2 171	2 487	2 449	2 300	2 387	2 474	2 667
2 791	2 904	2 737	2 849	2 723	2 613	2 950	2 825	2 717	2 593
2 703	2 836	2 938	2 975	3 064	3 092	3 063	2 991		

资料来源：Andrews & Herzberg (1985).

附录 1.18　　美国 1963 年 4 月—1971 年 7 月短期国库券的月度收益率序列

0.002 38	0.002 38	0.002 36	0.002 5	0.002 54	0.002 6	0.002 85	0.002 81
0.002 41	0.002 88	0.002 87	0.002 92	0.002 94	0.002 73	0.002 71	0.002 82
0.002 67	0.002 73	0.002 93	0.002 85	0.002 96	0.002 81	0.003 26	0.003 21
0.003 15	0.003 19	0.003 13	0.003 13	0.003 19	0.003 13	0.003 30	0.003 19
0.003 15	0.003 55	0.003 70	0.003 71	0.003 64	0.003 81	0.003 72	0.003 68
0.003 74	0.003 89	0.004 15	0.003 89	0.003 43	0.003 77	0.003 68	0.003 64
0.003 38	0.002 83	0.002 71	0.003 00	0.003 09	0.003 17	0.003 43	0.003 47
0.003 55	0.003 60	0.003 98	0.003 85	0.003 89	0.004 44	0.004 53	0.004 44
0.004 32	0.004 06	0.004 32	0.004 61	0.003 98	0.005 00	0.004 87	0.004 70
0.004 32	0.005 08	0.004 78	0.005 08	0.005 93	0.005 50	0.005 93	0.005 42
0.005 40	0.005 40	0.006 31	0.005 34	0.005 46	0.005 42	0.005 46	0.004 95
0.005 00	0.005 08	0.004 83	0.004 44	0.003 77	0.003 55	0.003 38	0.002 64
0.002 83	0.003 05	0.003 38	0.004 06				

资料来源：Hipel and Mcleod (1994).

附录 1.19　　　M1, U.S. 1959.1—1992.2

143.1	140.3	139.4	140.7	139.6	140.4	141.2	140.9	141.3	141.7	142.8	144.7
144.4	140.9	139.5	140.8	138.7	139.0	140.0	140.4	141.6	142.3	143.4	145.7
145.7	142.8	141.8	143.5	141.8	142.4	142.8	142.7	144.3	145.7	147.6	150.5
150.2	146.9	146.0	148.0	145.8	146.2	146.4	145.8	146.9	148.4	150.2	153.3
153.6	150.1	149.3	151.5	149.3	151.4	151.3	150.9	152.5	154.4	156.7	159.0
159.4	155.4	154.6	156.8	154.2	155.5	157.1	157.0	159.4	161.3	163.1	166.4
166.9	161.9	161.5	164.2	160.3	162.2	163.5	162.8	165.6	168.2	169.9	174.4
175.6	170.3	170.4	174.1	169.6	171.7	171.0	170.0	172.7	173.4	174.6	178.6
178.4	173.4	174.6	176.6	174.1	177.4	179.1	179.0	181.7	183.9	185.7	190.3
189.0	184.9	185.4	189.3	186.5	190.2	191.9	191.4	193.9	196.3	199.6	204.8
205.9	199.3	199.8	203.6	199.4	202.3	203.3	201.5	203.2	205.0	207.0	211.4
212.9	204.0	205.5	210.1	206.2	208.9	210.1	210.0	212.8	214.4	216.7	222.2
222.6	216.6	218.6	223.7	221.1	225.2	227.5	225.9	227.7	229.1	231.2	236.9
237.5	231.4	234.2	239.5	234.7	238.8	241.8	241.3	244.5	247.0	250.5	258.9
259.4	251.2	251.6	257.0	253.6	259.3	261.1	258.6	259.5	261.4	265.6	273.3
271.8	264.1	266.5	271.6	266.3	271.5	273.5	271.0	272.6	274.8	278.8	285.2
281.8	273.3	276.4	281.4	278.1	286.0	288.0	286.3	287.8	288.5	293.5	299.0
296.8	289.0	291.4	299.9	295.1	299.4	302.3	301.0	302.5	307.0	309.7	318.6
317.7	309.0	312.2	322.7	315.6	321.7	326.3	324.3	327.7	332.0	335.4	344.1
343.4	332.0	334.9	347.5	342.4	349.4	353.9	351.7	357.0	359.4	362.9	372.5
367.8	356.4	360.8	376.2	367.1	376.7	383.3	381.9	385.6	387.7	389.8	398.6
390.7	380.9	382.4	387.1	377.8	387.6	394.8	398.5	404.9	411.0	416.1	419.8
416.5	405.7	412.5	431.3	418.6	423.0	427.9	426.1	427.3	429.8	435.2	447.2
448.7	432.6	435.8	451.3	441.1	446.5	449.6	450.0	456.4	466.0	474.5	486.0
483.0	474.2	482.9	498.7	494.1	503.7	510.7	508.5	511.5	517.4	522.1	533.4
530.4	517.6	524.2	539.2	530.8	541.4	543.3	539.0	542.5	542.1	549.6	564.5
561.1	551.9	558.3	575.0	569.4	585.2	592.0	594.8	602.2	605.5	615.1	633.5
626.8	613.1	624.6	647.2	645.7	663.5	674.0	679.1	685.2	692.8	709.5	740.6
737.5	717.1	723.5	752.5	739.9	744.4	746.8	745.0	745.2	753.7	756.0	765.9
764.7	745.0	752.1	778.3	763.8	778.8	785.6	781.3	780.0	780.8	787.1	803.2
793.0	772.3	775.2	791.3	767.2	773.8	781.7	777.4	778.5	784.5	791.4	811.9
802.4	788.3	796.2	818.0	797.3	810.8	812.9	814.5	818.9	817.6	826.1	844.3
833.2	823.4	835.0	852.9	841.9	857.8	861.9	864.2	867.3	875.0	893.4	916.8
918.1	916.5										

数据来源：Hipel and Mcleod (1994).

附录 1.20　　**天然气炉输入—输出数据**

输出序列（在输出气体中 CO_2 的百分浓度）

53.8	53.6	53.5	53.5	53.4	53.1	52.7	52.4
52.2	52.0	52.0	52.4	53.0	54.0	54.9	56.0
56.8	56.8	56.4	55.7	55.0	54.3	53.2	52.3
51.6	51.2	50.8	50.5	50.0	49.2	48.4	47.9
47.6	47.5	47.5	47.6	48.1	49.0	50.0	51.1
51.8	51.9	51.7	51.2	50.0	48.3	47.0	45.8
45.6	46.0	46.9	47.8	48.2	48.3	47.9	47.2
47.2	48.1	49.4	50.6	51.5	51.6	51.2	50.5
50.1	49.8	49.6	49.4	49.3	49.2	49.3	49.7
50.3	51.3	52.8	54.4	56.0	56.9	57.5	57.3
56.6	56.0	55.4	55.4	56.4	57.2	58.0	58.4
58.4	58.1	57.7	57.0	56.0	54.7	53.2	52.1
51.6	51.0	50.5	50.4	51.0	51.8	52.4	53.0
53.4	53.6	53.7	53.8	53.8	53.8	53.3	53.0
52.9	53.4	54.6	56.4	58.0	59.4	60.2	60.0
59.4	58.4	57.6	56.9	56.4	56.0	55.7	55.3
55.0	54.4	53.7	52.8	51.6	50.6	49.4	48.8
48.5	48.7	49.2	49.8	50.4	50.7	50.9	50.7
50.5	50.4	50.2	50.4	51.2	52.3	53,2	53.9
54.1	54.0	53.6	53.2	53.0	52.8	52.3	51.9
51.6	51.6	51.4	51.2	50.7	50.0	49.4	49.3
49.7	50.6	51.8	53.0	54.0	55.3	55.9	55.9
54.6	53.5	52.4	52.1	52.3	53.0	53.8	54.6
55.4	55.9	55.9	55.2	54.4	53.7	53.6	53.6
53.2	52.5	52.0	51.4	51.0	50.9	52.4	53.5
55.6	58.0	59.5	60.0	60.4	60.5	60.2	59.7
59.0	57.6	56.4	55.2	54.5	54.1	54.1	54.4
55.5	56.2	57.0	57.3	57.4	57.0	56.4	55.9
55.5	55.3	55.2	55.4	56.0	56.5	57.1	57.3
56.8	55.6	55.0	54.1	54.3	55.3	56.4	57.2
57.8	58.3	58.6	58.8	58.8	58.6	58.0	57.4
57.0	56.4	56.3	56.4	56.4	56.0	55.2	54.0
53.0	52.0	51.6	51.6	51.1	50.4	50.0	50.0
52.0	54.0	55.1	54.5	52.8	51.4	50.8	51.2
52.0	52.8	53.8	54.5	54.9	54.9	54.8	54.4
53.7	53.3	52.8	52.6	52.6	53.0	54.3	56.0
57.0	58.0	58.6	58.5	58.3	57.8	57.3	57.0

输入序列（输入天然气速率（ft. / min.））

−0.109	0	0.178	0.339	0.373	0.441	0.461	0.348
0.127	−0.180	−0.588	−1.055	−1.421	−1.520	−1.302	−0.814
−0.475	−0.193	0.088	0.435	0.771	0.866	0.875	0.891
0.987	1.263	1.775	1.976	1.934	1.866	1.832	1.767
1.608	1.265	0.790	0.360	0.115	0.088	0.331	0.645
0.960	1.409	2.670	2.834	2.812	2.483	1.929	1.485
1.214	1.239	1.608	1.905	2.023	1.815	0.535	0.122
0.009	0.164	0.671	1.019	1.146	1.155	1.112	1.121
1.223	1.257	1.157	0.913	0.620	0.255	−0.280	−1.080
−1.551	−1.799	−1.825	−1.456	−0.944	−0.570	−0.431	−0.577
−0.960	−1.616	−1.875	−1.891	−1.746	−1.474	−1.201	−0.927
−0.524	0.040	0.788	0.943	0.930	1.006	1.137	1.198
1.054	0.595	−0.080	−0.314	−0.288	−0.153	−0.109	−0.187
−0.255	−0.229	−0.007	0.254	0.330	0.102	−0.423	−1.139
−2.275	−2.594	−2.716	−2.510	−1.790	−1.346	−1.081	−0.910
−0.876	−0.885	−0.800	−0.544	−0.416	−0.271	0	0.403
0.841	1.285	1.607	1.746	1.683	1.485	0.993	0.648
0.577	0.577	0.632	0.747	0.900	0.993	0.968	0.790
0.399	−0.161	−0.553	−0.603	−0.424	−0.194	−0.049	0.060
0.161	0.301	0.517	0.566	0.560	0.573	0.592	0.671
0.933	1.337	1.460	1.353	0.772	0.218	−0.237	−0.714
−1.099	−1.269	−1.175	−0.676	0.033	0.556	0.643	0.484
0.109	−0.310	−0.697	−1.047	−1.218	−1.183	−0.873	−0.336
0.063	0.084	0	0.001	0.209	0.556	0.782	0.858
0.918	0.862	0.416	−0.336	−0.959	−1.813	−2.378	−2.499
−2.473	−2.330	−2.053	−1.739	−1.261	−0.569	−0.137	−0.024
−0.050	−0.135	−0.276	−0.534	−0.871	−1.243	−1.439	−1.422
−1.175	−0.813	−0.634	−0.582	−0.625	−0.713	−0.848	−1.039
−1.346	−1.628	−1.619	−1.149	−0.488	0.160	−0.007	−0.092
−0.620	−1.086	−1.525	−1.858	−2.029	−2.024	−1.961	−1.952
−1.794	−1.302	−1.030	−0.918	−0.798	−0.867	−1.047	−1.123
−0.876	−0.395	0.185	0.662	0.709	0.605	0.501	0.603
0.943	1.223	1.249	0.824	0.102	0.025	0.382	0.922
1.032	0.866	0.527	0.093	−0.458	−0.748	−0.947	−1.029
−0.928	−0.645	−0.424	−0.276	−0.158	−0.033	0.102	0.251
0.280	0	−0.493	−0.759	−0.824	−0.740	−0.528	−0.204
0.034	0.204	0.253	0.195	0.131	0.017	−0.182	−0.262

数据来源：Cox& Jenkins. *Time Series Analysis*: *Forecasting and Control* 数据集 J. 1976.

附录 1.21　1978—2002 年中国农村居民家庭人均纯收入序列 $\{x_t\}$，生活消费支出序列 $\{y_t\}$ 及对数纯收入序列 $\{\ln x_t\}$，对数生活消费支出序列 $\{\ln y_t\}$

年份	纯收入		生活消费支出	
	x_t	$\ln x_t$	y_t	$\ln y_t$
1978	133.6	4.894 85	116.1	4.754 45
1979	160.7	5.079 54	134.5	4.901 56
1980	191.3	5.253 84	162.2	5.088 83
1981	223.4	5.408 96	190.8	5.251 23
1982	270.1	5.598 79	220.2	5.394 54
1983	309.8	5.735 93	248.3	5.514 64
1984	355.3	5.872 96	273.8	5.612 40
1985	397.6	5.985 45	317.4	5.760 16
1986	423.8	6.049 26	357.0	5.877 74
1987	462.6	6.136 86	398.3	5.987 21
1988	544.9	6.300 60	476.7	6.166 89
1989	601.5	6.399 43	535.4	6.283 01
1990	686.3	6.531 31	584.6	6.370 93
1991	708.6	6.563 29	619.8	6.429 40
1992	784.0	6.664 41	659.8	6.491 94
1993	921.6	6.826 11	769.7	6.646 00
1994	1 221.0	7.107 43	1 016.8	6.924 42
1995	1 577.7	7.363 72	1 310.4	7.178 09
1996	1 926.1	7.563 25	1 572.1	7.360 17
1997	2 090.1	7.644 97	1 617.2	7.388 45
1998	2 162.0	7.678 79	1 590.3	7.371 68
1999	2 210.3	7.700 88	1 577.4	7.363 53
2000	2 253.4	7.720 20	1 670.1	7.420 64
2001	2 366.4	7.769 13	1 741.0	7.462 21
2002	2 476.0	7.814 40	1 834.0	7.514 25

附录 1.22　1880—1985 年全球气表平均温度改变值序列　单位：摄氏度

−.40	−.37	−.43	−.47	−.72	−.54	−.47	−.54	−.39	−.19
−.40	−.44	−.44	−.49	−.38	−.41	−.27	−.18	−.38	−.22
−.03	−.09	−.28	−.36	−.49	−.25	−.17	−.45	−.32	−.33
−.32	−.29	−.32	−.25	−.05	−.01	−.26	−.48	−.37	−.20
−.15	−.08	−.14	−.13	−.12	−.10	.13	−.01	.06	−.17
−.01	.09	.05	−.16	.05	−.02	.04	.17	.19	.05

续前表

.15	.13	.09	.04	.11	−.03	.03	.15	.04	−.02
−.13	.02	.07	.20	−.03	−.07	−.19	.09	.11	.06
.01	.08	.02	.02	−.27	−.18	−.09	−.02	−.13	.02
.03	−.12	−.08	.17	−.09	−.04	−.24	−.16	−.09	.12
.27	.42	.02	.30	.09	.05				

注：平均温度为零点。

数据来源：James Hansen and Sergej Lebedeff (1987).

附录 2

附录 2.1 字符型数据的常用输入格式

输入格式	描述	w 的范围	w 的缺省值
$ w.	输入标准字符数据	1～200	
$ CHARw.	输入含有空格的字符数据	1～200	1 或变量长度
$ QUOTEw.	从数据值中移走引号	1～200	8
$ UPCASEw.	将所有的字符转换为大写读入	1～200	8 或变量长度

附录 2.2 字符型数据的常用输出格式

输入格式	描述	w 的范围	w 的缺省值
$ w.	输出标准字符数据	1～200	1 或变量长度
$ CHARw.	输出标准字符数据	1～200	1 或变量长度
$ QUOTEw.	输出包含引号的数据值	1～200	8 或变量长度
$ UPCASEw.	将所有的字符转换为大写输出	1～200	8 或变量长度

附录 2.3 数值型数据的常用输入格式

输入格式	描述	w 的范围	w 的缺省值
$w.d$	输入总长度为 w，d 位小数的标准数值数据	1～32	—
COMMA$w.d$	移走数值中嵌入的逗号、小数点和括号	1～32	1
E$w.d$	读取用科学记数法表示的数据	7～32	12
PERCENT$w.$	转换百分位数为数值	1～32	1

附录 2.4 数值型数据的常用输出格式

输入格式	描述	w 的范围	w 的缺省值
$w.d$	输出总长度为 w，d 位小数的标准数值数据	1～32	—
BEST$w.$	选择最好的表示法	1～32	1
COMMA$w.d$	用含有逗号、小数点的格式输出数据	2～32	6
DOLLAR$w.d$	用含有美元符号、逗号及小数点的格式输出数据	1～32	1
E$w.d$	用科学记数法输出数据	7～32	12
PERCENT$w.$	以百分位数的形式输出数据	4～32	6

附录 2.5 时间数据的常用输入格式

输入格式	描述	可读形式举例	w 的范围	w 的缺省值
DATE$w.$	以日-月-年的顺序输入日期值(ddmmmyy)	08jan05 8jan2005 8-jan-2005	7～32	7
DDMMYY$w.$	以日-月-年的顺序输入日期值(ddmmyy)	080105 08/01/05 08-01-05	6～32	6
MMDDYY$w.$	以月-日-年的顺序输入日期值(mmddyy)	010805 01/08/05 01-08-05	6～32	6
YYMMDD$w.$	以年-月-日的顺序输入日期值(mmddyy)	050108 05/01/08 05-01-08	6～32	6
MONYY$w.$	以月-年顺序输入月份值(mmmyy)	Jan05 Jan2005	5～32	5
YYQ$w.$	以年-季节的顺序输入季节值	05q1 2005q1	4～32	4
TIME$w.$	以小时-分-秒形式输入时间值	5：3：2 23：42：12	5～32	8
DATETIME$w.$	以日-月-年-小时-分-秒形式输入日期与时间值	08jan2005/5：3：2 08jan05/23：42：12	13～32	18

附录 2.6 **时间数据的常用输出格式**

输入格式	描述	可读形式举例	*w* 的范围	*w* 的缺省值
DATE*w*.	以日-月-年的顺序输出日期值（ddmmmyy）	08jan05 8jan2005	5～9	7
DDMMYY*w*.	以日-月-年的顺序输出日期值（ddmmyy）	08/01/05	2～10	8
MMDDYY*w*.	以月-日-年的顺序输出日期值（mmddyy）	01/08/05	2～10	8
YYMMDD*w*.	以年-月-日的顺序输出日期值（mmddyy）	05-01-08	2～10	8
WEEKDATE*w*.	输出星期与日期值	Sat，jan 08，2005	3～37	29
DAY*w*.	只输出日期值	08	2～32	2
MONYY*w*.	以月-年顺序输出月份值	Jan05 Jan2005	5～7	7
YYMON*w*.	以年-月顺序输出月份值	05 Jan 2005Jan	5～7	7
MONTH*w*.	只输出月份值	01 jan	2～32	2
YYQ*w*.	以年-季节的顺序输出季节值	2005q1	4～32	6
QTR*w*.	只输出季节值	1	1～32	1
TIME*w*.	以小时-分-秒形式输出时间值	5：3：2 23：42：12	2～20	8
DATETIME*w*.	以日-月-年-小时-分-秒形式输出日期与时间值	08jan2005/5：3：2 08jan05/23：42：12	7～40	16

附录 3

X-11 过程输出表格说明

表格	说明
A1	原始序列
A2	先验月度调整因子
A3	进行完先验月度因子调整的原始序列
A4	先验交易日调整
A5	先验调整过的或原始的序列
A13	ARIMA 预报
A14	ARIMA 向后外推
A15	先验调整过的或原始的序列被 ARIMA 预报、反向外推的结果
B1	先验调整后的或原始的序列
B2	趋势起伏
B3	未改动的季节、不规则（S-I）比
B4	对极端 S-I 比的替换值
B5	季节因子
B6	季节调整后的序列
B7	趋势起伏

续前表

表格	说明
B8	未改动的 S-I 比
B9	极端 S-I 比的替换值
B10	季节因子
B11	季节调整后的序列
B13	不规则序列
B14	从交易日回归中排除的极端不规则值
B15	初始交易日回归
B16	交易日调整因子
B17	不规则成分的初始权重
B18	由合并星期权重导出的交易日因子
B19	经过交易日和先验变化调整后的原始序列
C1	原始序列被初始权重修改后再对交易日和先验变化调整的结果
C2	趋势起伏
C4	修改后的 S-I 比
C5	季节因子
C6	季节调整后的序列
C7	趋势起伏
C9	修改后的 S-I 比
C10	季节因子
C11	季节调整后的序列
C13	不规则序列
C14	从交易日回归中排除的极端不规则值
C15	最终的交易日回归
C16	由回归系数导出的最终的交易日调整因子
C17	不规则成分的最终的权重
C18	由合并星期权重导出的最终交易日因子
C19	经过交易日和先验变动调整的原始序列
D1	用最终的权重进行修改并进行了交易日和先验变动调整的原始序列
D2	趋势起伏
D4	修改的 S I 比
D5	季节因子
D6	季节调整后的序列
D7	趋势起伏
D8	最终的未修改 S-I 比
D9	最终的极端 S-I 比的替换值

续前表

表格	说明
D10	最终的季节因子
D11	最终的季节调整后的序列
D12	最终的趋势起伏
D13	最终的不规则序列
E1	替换了异常值的原始序列
E2	修改了的季节调整后的序列
E3	修改了的不规则序列
E4	年度总计的比例
E5	原始序列的修改百分比
E6	最终的季节调整后序列的改变百分比
F1	MCD 滑动平均
F2	概况性度量
G1	最终季节调整后序列和趋势
G2	带极端值的 S-I 比图表，无极端值的 S-I 比图表及最终的季节因子图表
G3	按日历次序的带极端值的 S-I 比，无极端值的 S-I 比及最终季节因子的图表
G4	最终的不规则列和最终修改的不规则列的图表

参考文献

1. George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel 著，顾岚译. 时间序列分析预测与控制(第三版). 北京：中国统计出版社，1997

2. [英]C. 查特菲尔德著，骆振华译. 时间序列分析引论(第二版). 厦门：厦门大学出版社，1987

3. [英]特伦斯·C·米尔斯著，俞卓菁译. 金融时间序列的经济计量学模型(第二版). 北京：经济科学出版社，2002

4. 高惠璇等编译. SAS 系统 SAS/ETS 软件使用手册. 北京：中国统计出版社，1998

5. Dickey, D. A and Fuller, W. A, 'Distribution of the Estimators for Autoregressive Time Series with a unit Root', Journal of the American Statistical Associateion, 1979

6. Engle, R. F, "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U. K. Inflation", *Econometrics*, 1982

7. Granger, C. W. J and Joyeux, R. , "An Introduction to Long Memory Time Series Models and Fractional Differencing", *Journal of Time Series Analysis*, 1980

8. Granger, C. W. J and Swanson, N. , "Future Development in the Study of Cointegrated Variables", *Oxford Bulletin of Economics and Statistics*, 1996